Cerebral Cortex, August 2020;30: 4515–4527

doi: 10.1093/cercor/bhaa054
Advance Access Publication Date: 7 March 2020
Original Article

ORIGINAL ARTICLE

Top–Down Inhibitory Mechanisms Underlying Auditory–Motor Integration for Voice Control: Evidence by TMS

Dongxu Liu[1], Guangyan Dai[1], Churong Liu[2], Zhiqiang Guo[3], Zhiqin Xu[1], Jeffery A. Jones[4], Peng Liu[1] and Hanjun Liu[1,5]

[1]Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, China, [2]Rehabilitation Training Center, Guangzhou 999 Brain Hospital, Guangzhou 510510, China, [3]Department of Computer Science and Technology, Zhuhai College of Jilin University, Zhuhai 519041, China, [4]Psychology Department and Laurier Centre for Cognitive Neuroscience, Wilfrid Laurier University, Waterloo, Ontario, Canada and [5]Guangdong Provincial Key Laboratory of Brain Function and Disease, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China

Address correspondence to Hanjun Liu, PhD, Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, China. Email: lhanjun@mail.sysu.edu.cn
Dongxu Liu and Guangyan Dai have contributed equally to this work

Abstract

The dorsolateral prefrontal cortex (DLPFC) has been implicated in auditory–motor integration for accurate control of vocal production, but its precise role in this feedback-based process remains largely unknown. To this end, the present event-related potential study applied a transcranial magnetic stimulation (TMS) protocol, continuous theta-burst stimulation (c-TBS), to disrupt cortical activity in the left DLPFC as young adults vocalized vowel sounds while hearing their voice unexpectedly shifted upwards in pitch. The results showed that, as compared to the sham condition, c-TBS over left DLPFC led to significantly larger vocal compensations for pitch perturbations that were accompanied by significantly smaller cortical P2 responses. Source localization analyses revealed that this brain activity pattern was the result of reduced activation in the left superior frontal gyrus and right inferior parietal lobule (supramarginal gyrus). These findings demonstrate c-TBS-induced modulatory effects of DLPFC on the neurobehavioral processing of vocal pitch regulation, suggesting that disrupting prefrontal function may impair top–down inhibitory control mechanisms that prevent speech production from being excessively influenced by auditory feedback, resulting in enhanced vocal compensations for feedback perturbations. This is the first study that provides direct evidence for a causal role of the left DLPFC in auditory feedback control of vocal production.

Key words: auditory feedback, continuous theta-burst stimulation, dorsolateral prefrontal cortex, speech motor control, transcranial magnetic stimulation

Introduction

Auditory feedback plays an important role in sensorimotor integration for accurate control of speech production. One important aspect of speech motor control that involves auditory feedback is represented by the rapid compensatory vocal responses that are elicited by unexpected perturbations in voice fundamental frequency (F_0), intensity, and formant frequencies (e.g., F1) (Burnett et al. 1998; Houde and Jordan 1998; Bauer et al. 2006; Liu and Larson 2007). These compensatory responses correct for perceived errors in auditory feedback and can be modulated by task demands (Natke et al. 2003; Chen et al. 2007), shaped by learning experience (Liu et al. 2010; Behroozmand et al. 2014), and affected by neurological diseases (Huang et al. 2016; Parrell et al. 2017; Ranasinghe et al. 2017). It is hypothesized that the speech motor control system monitors the mismatches between the intended and actual vocal output and initiates corrective motor commands to ensure accurate speech production (Guenther et al. 2006; Golfinopoulos et al. 2010; Hickok et al. 2011; Houde and Nagarajan 2011; Guenther and Vladusich 2012). The neural mechanisms underlying the role of auditory feedback in speech motor control, however, are still far from clear.

There is accumulating evidence suggesting the involvement of higher-level cognitive processes that drive compensatory adjustment of vocal motor behavior in response to auditory feedback errors. For example, focused attention on pitch perturbations in voice auditory feedback has been shown to elicit enhanced vocal compensations and event-related potential (ERP) P2 responses, while divided attention has been shown to elicit the opposite pattern of effects (Tumber et al. 2014; Hu et al. 2015; Liu et al. 2015). Increases in the load level of divided attention have been shown to lead to larger vocal compensations for pitch perturbations, larger N1 responses, and smaller P2 responses (Liu et al. 2018). On the other hand, training-induced improvements in working memory capacity were significantly correlated with decreased vocal compensations for pitch perturbations that were accompanied by increased cortical P2 responses that were source-localized in the inferior frontal gyrus (IFG), inferior parietal lobule (IPL), and insula (Guo et al. 2017). Consistently, Ranasinghe et al. (2017) found that patients with Alzheimer's disease (AD) exhibited enhanced magnitudes and reduced durations of vocal compensations for pitch perturbations that were predicted by their executive and memory dysfunctions when compared to healthy controls. Together, these findings suggest an interplay between auditory–motor integration for vocal pitch regulation and higher-level cognitive functions such as attention, working memory, and executive control and that impairment in these cognitive abilities may compromise auditory–vocal integration as reflected by enhanced vocal compensations for pitch perturbations. Specifically, it has been suggested that top–down modulatory mechanisms may exist to inhibit compensatory vocal responses to feedback perturbations (Guo et al. 2017; Ranasinghe et al. 2017). However, the neural correlates of this top–down control of vocal production remain largely unknown.

Given the critical role of the dorsolateral prefrontal cortex (DLPFC) in working memory (Edin et al. 2009), attentional control (Brosnan and Wiegand 2017), and executive functions (Mansouri et al. 2009), the DLPFC may be a core region that mediates top–down influences on auditory feedback control of vocal production. This idea is supported by a recent magnetoencephalograph (MEG) study by Ranasinghe et al. (2019), where patients with AD exhibited less brain activity in the left DLPFC and larger vocal compensations than those observed in healthy controls. Moreover, lower brain activity in the left DLPFC was predictive of larger vocal compensations across both the AD and control groups. Similarly, in addition to the activation observed in the superior temporal gyrus (STG), premotor cortex (PMC), supplementary motor area (SMA), IFG, and IPL (particularly the supramarginal gyrus, SMG) when participants compensate for voice feedback perturbations (Zarate and Zatorre 2008; Parkinson et al. 2012; Chang et al. 2013; Behroozmand et al. 2015; Guo et al. 2016; Kort et al. 2016), the left DLPFC has been found to be active when non-singers are asked to ignore or compensate for pitch perturbations during singing (Zarate and Zatorre 2008). As well, patients with Parkinson's disease (PD) who received extensive voice treatment exhibited increased activity in the right DLPFC that was significantly correlated with the treatment outcome (Narayana et al. 2010). Related work has shown that the DLPFC receives auditory afferents from the caudal auditory belt region (Romanski et al. 1999) and exerts top–down inhibitory regulation of auditory processing (Knight et al. 1989; Mitchell et al. 2005). Therefore, it is plausible to hypothesize that the DLPFC exerts top–down influences on auditory feedback control of vocal production. There is no direct, causal evidence, however, to support this hypothesis.

To address this important question, we used the frequency-altered feedback (FAF) paradigm and combined it with a continuous theta-burst stimulation protocol (c-TBS). TBS is a form of repetitive transcranial magnetic stimulation (rTMS) protocol that induces strong and long-lasting effects on cortical excitability. Specifically, c-TBS disrupts the cortical activity for up to 60 min after less than 1-min stimulation, while intermittent TBS (i-TBS) facilitates the cortical activity for up to 20 min after less than 4-min stimulation (Huang et al. 2005). TMS has been used to explore the causal relationship between brain regions and speech/language functions (Devlin and Watkins 2007; Murakami et al. 2013) and to treat neurogenic speech–language disorders (Murdoch and Barwood 2013). Recently, a number of TMS studies have been performed to examine the contributions of specific cortical regions to speech production (Shum et al. 2011; Restle et al. 2012; Lega et al. 2016; Finkel et al. 2019). For example, low-frequency rTMS over left IPL (SMG) led to diminished adaptive responses to F1 feedback perturbations during the syllable production, indicating that the IPL (SMG) plays a central role in sensorimotor adaptation to speech production (Shum et al. 2011). Finkel et al. (2019) found that i-TBS over right somatosensory larynx cortex (S1) enhanced participants' pitch accuracy and stability when they performed a pitch-matching singing task with noise-masked auditory feedback, indicating a causal involvement of larynx-S1 in vocal pitch regulation. The causal relationship between the DLPFC and auditory feedback control of vocal production, however, has never been investigated.

Based on the previous findings that showed activation of the left DLFPC in compensating for vocal pitch perturbations (Zarate and Zatorre 2008) and a significant correlation between reduced activity in the left DLPFC and enhanced vocal compensations (Ranasinghe et al. 2019), we used c-TBS to disrupt the cortical activity in the left DLPFC and examined the inhibitory effects of this stimulation on the vocal and ERP (N1 and P2) responses to pitch feedback perturbations. The N1 and P2 components have been found to be pronounced in response to pitch perturbations during vocal production (Behroozmand et al. 2009; Liu et al. 2011; Scheerer et al. 2013). These ERP components are thought to

reflect the early detection of a mismatch between intended and actual auditory feedback, as well as later interactions between the auditory and motor systems; both ERPs are associated with higher cognitive aspects of vocal output monitoring (Korzyukov et al. 2012a; Behroozmand et al. 2014; Guo et al. 2017). In light of previous results that showed a negative relationship between the magnitudes of vocal compensations and the amplitudes of P2 responses in the fronto-tempo-parietal regions (Guo et al. 2017) and brain activity in the left DLPFC (Ranasinghe et al. 2019), we expected that c-TBS over left DLPFC would lead to larger vocal responses and smaller cortical responses to pitch perturbations, supporting the existence of prefrontally mediated top–down mechanisms that exert inhibitory influences on auditory–motor control of vocal production.

Methods

Subjects

Twenty-six college students (13 females and 13 males; mean age: 22.85 ± 1.35 years) from Sun Yat-sen University of China participated in the present study. All participants were right-handed native-Mandarin speakers. Participants were included in the experiment if they met the following inclusion criteria: no pregnancy; no intake of neuropsychiatric drugs; no implanted medical device; no experience of musical training; and no history of speech, language, hearing, and neurological disorders. One participant did not complete the experiment because she felt uncomfortable while receiving c-TBS. Four other participants were also excluded from the final data pool because their electroencephalograph (EEG) data did not fit the inclusion criteria (see below). Therefore, data from 21 participants (10 females and 11 males, mean age = 22.71 ± 1.41 years) were retained in the final statistical analyses. Written informed consent was obtained from all participants and the research protocol (No. 2017-254) was approved by the Institutional Review Board of The First Affiliated Hospital of Sun Yat-sen University in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Transcranial Magnetic Stimulation

Magnetic stimulation was performed using a CCY-IA TMS instrument (YIRUIDE Co.) equipped with a 7-cm outer-diameter figure-of-eight coil. The resting motor threshold (RMT) was determined as the minimal stimulation intensity sufficient to induce a motor-evoked potential (MEP) in the right first dorsal interosseous muscle of at least 50 μV in at least 5 out of 10 TMS pulses on the contralateral M1 (Grossheinrich et al. 2009). The scalp location for stimulating the left DLPFC was identified as the F3 position located on a commercial standard extended 10-20 EEG caps (YIRUIDE Co.) that was designed according to the International 10-20 system (Herwig et al. 2003). In the International 10-20 system, adjacent electrodes are either 10% or 20% of the total distance between the nasion and inion, or left and right preauricular points, away from each other. The size of the EEG cap was appropriately chosen to fit each participant's head based on their individual head size. The coil was placed closely tangential to the surface of the skull during the real c-TBS, while it was placed at 90° to the surface of the skull when the sham c-TBS was performed. A standard c-TBS protocol consisted of bursts of three pulses at 50 Hz (20 ms between each pulse) that were repeatedly presented at 200-ms interval over 40 s for a total of 600 pulses (Huang et al. 2005). Active c-TBS protocol was applied at 80% RMT intensity (Grossheinrich et al. 2009; Teng et al. 2017). All participants received both real and sham c-TBS on separate days at least 7 days apart. The order of the two stimulation tasks was counterbalanced across all participants.

Experimental Design

It has been demonstrated that c-TBS decreases cortical excitability for up to 60 min (Hoogendam et al. 2010). In order to maximize the stimulation after-effects, the FAF-based vocal production experiment began immediately after the c-TBS task and completed within 25 min. All participants were instructed to start vocalizing the vowel sound /u/ at their habitual vocal pitch and loudness when a black cross appeared on the computer monitor and to maintain their vocalizations as steady as they could until the black cross disappeared. During each 6-s long vocalization, the participants' voice was pseudorandomly shifted upward in pitch by either 200 or 500 cents (100 cents = 1 semitone) for a duration of 200 ms. Given the modulatory effects of perturbation magnitude on the vocal and ERP responses (Behroozmand et al. 2009; Liu et al. 2011; Scheerer et al. 2013), the magnitude of the pitch perturbations was manipulated to examine whether c-TBS-induced effects were stimulus-specific. Also, the two magnitudes of the pitch perturbations were pseudorandomly manipulated to reduce the ability for participants to adapt to the repetitive stimulus presentation (Behroozmand et al. 2012; Korzyukov et al. 2012b). The first pitch perturbation occurred 500–1000 ms after the onset of each vocalization, while the succeeding perturbations were presented with an interstimulus interval of 700–900 ms. Prior to initiating the next vocalization, participants were required to take a 5-s break to avoid vocal fatigue. Each participant produced 40 vocalizations, leading to a total of 200 trials that included 100 trials with +200 cents perturbations and 100 trials with +500 cents perturbations. An additional vocal production experiment was conducted with the same experimental parameters after participants received the sham stimulation.

Data Acquisition

Throughout the vocal production experiment, all participants sat on a comfortable chair in a sound-attenuated booth. An acoustic calibration procedure was performed by amplifying the intensity of the voice feedback heard by the participants so that it was 10 dB SPL higher than their voice output to minimize the masking effects of air-born and bone-conducted feedback. The voice signals were picked up by a dynamic microphone (DM2200, Takstar Inc.), amplified by a MOTU Ultralite Mk3 Firewire audio interface, and pitch-shifted by an Eventide Eclipse Harmonizer. A custom-developed MIDI software program (Max/MSP, v.5.0 by Cycling 74) was used to control the harmonizer to pitch-shift the voice signals. This program also generated the transistor–transistor logic (TTL) control pulses that marked the perturbation onset and the visual cues that guided the participants to start and stop the vocalizations. Amplified by an ICON NeoAmp headphone amplifier, the pitch-shifted voice signals were finally played back to the participants through insert earphones (ER-1, Etymotic Research Inc.). The original and feedback voice signals as well as the TTL control pulses were digitized at 10 kHz by a PowerLab A/D converter (ML880, AD Instruments) and recorded onto an iMac computer using LabChart software (v.7.0, AD Instruments).

The EEG signals were recorded from each participant's scalp using a 64-electrode Geodesic Sensor Net that was connected to a Net Amps 300 amplifier (Electrical Geodesics Inc.). The electrooculogram was monitored for artifacts with four electrodes placed above and at the outer canthus. Since this amplifier accepts scalp-electrode impedances up to 40–60 kΩ (Ferree et al. 2001), the impedance levels of individual sensors were carefully adjusted and kept below 50 kΩ throughout the EEG recording. In order to synchronize the voice and EEG signals, the TTL control pulses that marked the onset of each pitch perturbation were sent to the EEG recording system via a DIN synch cable. The EEG signals across all channels were referenced to the vertex (Cz) (Ferree et al. 2001) and recorded at 1 kHz using NetStation software (v. 4.5, Electrical Geodesics Inc.).

Vocal and EEG Data Analysis

A custom-developed IGOR PRO software program (v.6.0, Wavemetrics Inc.) was used to measure the magnitude and latency of vocal compensations for pitch perturbations. First, the voice F_0 contour in Hertz was extracted from the voice signals using Praat software (Boersma 2001) and converted to cents scale with the following formula: $\text{cents} = 100 \times (12 \times \log_2[F_0/\text{reference}])$ (reference = 195.997 Hz [G3 note]). Then, the voice F_0 contour in cents was segmented into epochs ranging from 200 ms before to 700 ms after the perturbation onset. All individual trials were visually inspected to reject those bad trials that were corrupted by vocal interruptions or signal processing errors. On average, 12% of the individual trials were rejected from further analysis. A presorting procedure was used to categorize the individual trials as compensatory or following responses based on the direction of the response to the perturbation on a trial-by-trial basis. To investigate the effects of c-TBS on the compensatory mechanisms of vocal motor control, only those compensatory trials that opposed the direction of the pitch perturbations were retained in the final averaging analysis. Overall, 56% of the individual trials were regarded as compensatory responses, which is in line with previous studies (Behroozmand et al. 2012; Li et al. 2013; Guo et al. 2016; Patel et al. 2019). The individual compensatory trials that contained no artifacts were normalized by subtracting the F_0 values after the onset of the pitch perturbation by the mean F_0 values in the baseline period (−200 to 0 ms) and averaged to generate an overall compensatory response for each condition. The magnitude and latency of one compensatory vocal response were defined as the peak F_0 value in cents and time in ms when the voice F_0 contours reached the minimum value.

The EEG data were analyzed offline using NetStation software. First, the EEG signals were band-pass filtered with cutoff frequencies of 1 and 20 Hz and segmented into epochs using a window ranging from −200 to +500 ms relative to the onset of the pitch perturbation. All segmented individual trials were then submitted to an artifact detection procedure, whereby any trials that exceeded ±55 μv of the moving average over an 80-ms window or contained more than 10 bad channels were excluded from further analysis. And files were marked bad and discarded if they contained artifacts in more than 20% of the epochs. An additional visual trial-by-trial inspection was conducted to ensure that all individual trials were appropriately rejected. On average, 86% of the trials were retained to produce the ERPs. Finally, all artifact-free trials were re-referenced to the average of electrodes on each mastoid and submitted to an averaging procedure, whereby they were averaged and baseline-corrected to generate an overall ERP response. Based on the prominent cortical responses to pitch perturbations observed in the two stimulation conditions, 24 electrodes in three regions of interest (ROI) were chosen for analysis: frontal area, including AF3, AFz, AF4, F5, F3, F1, Fz, F2, F4, F6; fronto-central area, including FC5, FC3, FC1, FCz, FC2, FC4, FC6; and central area, including C5, C3, C1, Cz, C2, C4, C6. The amplitudes and latencies of N1 and P2 components were measured from the averaged ERP response for each ROI as the negative and positive peak values and times in the time windows of 80–180 ms and 160–280 ms after the perturbation onset, respectively.

Source Localization

In order to examine c-TBS-induced changes in the neural correlates of auditory feedback control of vocal production, standardized low-resolution brain electromagnetic tomography (sLORETA) (Fuchs et al. 2002) in EEGLAB software (http://www.sccn.ucsd.edu/eeglab) was used to localize the neural sources of the N1 and P2 responses to pitch perturbations that differed between the c-TBS and sham conditions. The localization accuracy of sLORETA has been validated in studies that combined EEG with fMRI (Mulert et al. 2004) and intracerebral recordings (Zumsteg et al. 2006). For sLORETA, the intracerebral volume is partitioned in 6239 cortical gray matter voxels at 5-mm spatial resolution and the standardized current density is calculated in a realistic standardized head model (Fuchs et al. 2002) within the MNI152 template (Mazziotta et al. 2001). In the present study, the voxel-based sLORETA images were computed based on the averaged ERPs at the 5-ms time window of maximal global field power peaks in the N1 and P2 time windows and compared between the c-TBS and sham conditions using sLORETA-built-in voxel-wise randomization tests with 10 000 permutations. Multiple whole-brain comparisons were corrected based on statistical nonparametric mapping. The voxels with significant differences (for corrected $P < 0.05$) were specified in MNI coordinates and Brodmann areas (BA) that were determined using the EEGLAB software (Delorme and Makeig 2004). The statistical source results were superimposed on an anatomical template provided in BrainNet Viewer (Xia et al. 2013).

Statistical Analyses

The values of vocal and ERP responses were verified to be normally distributed using Kolmogorov–Smirnov test and thus analyzed using repeated measures analysis of variances (RM-ANOVAs) in SPSS (v.20.0). The magnitudes and latencies of vocal responses were subjected to two-way RM-ANOVAs, including two within-subject factors of perturbation magnitude (+200 vs. +500 cents) and stimulation condition (c-TBS vs. sham). The amplitudes and latencies of the N1 and P2 responses were subjected to three-way RM-ANOVAs, including three within-subject factors of perturbation magnitude, stimulation condition, and electrode site (frontal, fronto-central, and central). Subsidiary RM-ANOVAs were conducted if there were any significant higher-order interactions among any of those variables. Post hoc analyses for multiple comparisons were performed using Bonferroni correction. Probability values were corrected using Greenhouse–Geisser correction for multiple degrees of freedom when the assumption of Mauchly's test of sphericity for homogeneity of variance was violated. Effect sizes indexed by partial η^2 were calculated to quantify the proportion of

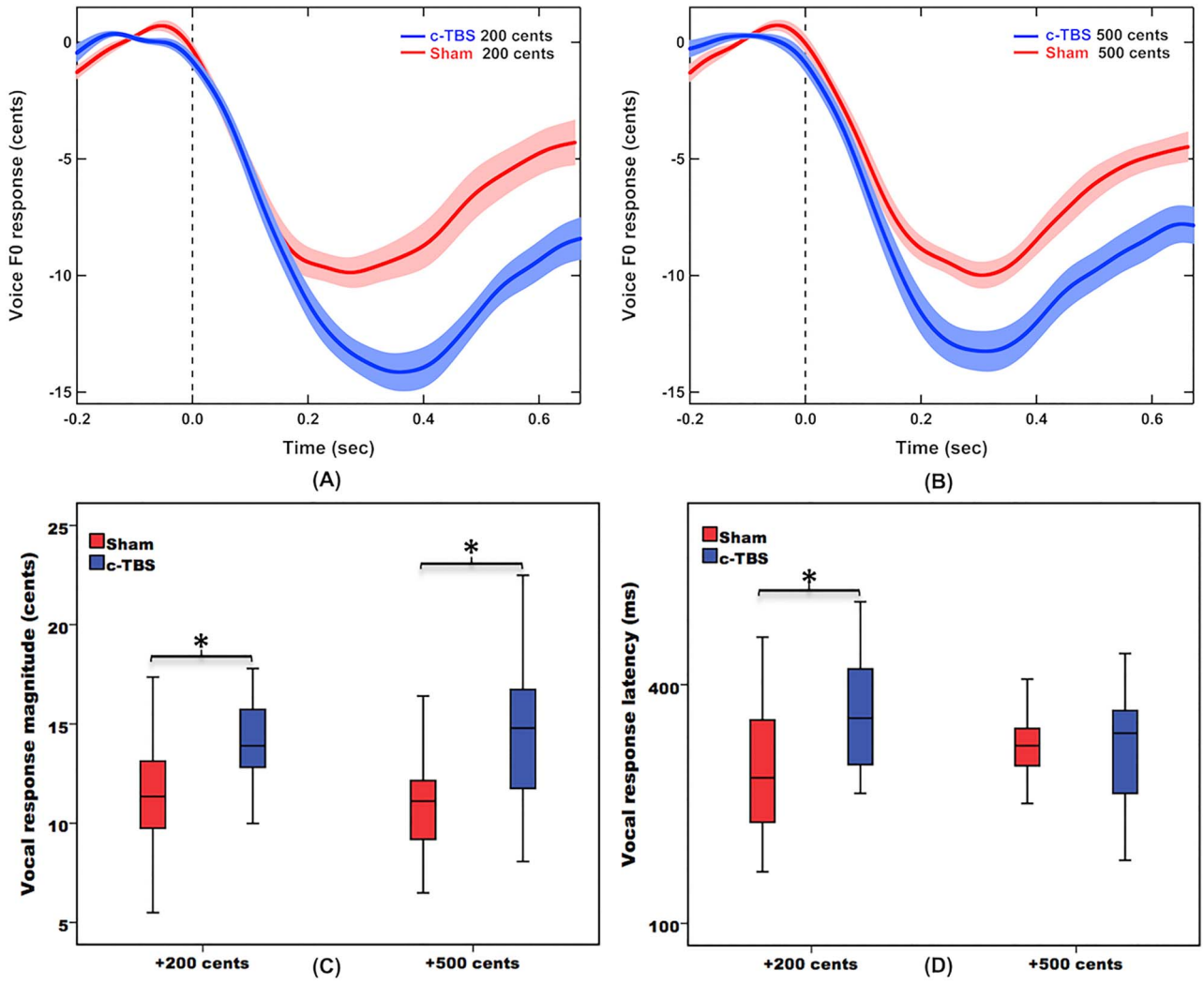

Figure 1. Top: Grand-averaged voice F_0 contours in response to +200 (*A*) and +500 cents (*B*) perturbations for the c-TBS (blue solid lines) and sham (red solid lines) conditions. Highlighted areas indicate the standard errors of the mean vocal responses. Vertical dashed lines indicate the onset of the pitch perturbation. Bottom: Box plots of the magnitudes (*C*) and latencies (*D*) of vocal responses to +200 and +500 cents perturbations during the c-TBS (blue) and sham (red) conditions. The asterisks indicate significant differences between conditions. The top and bottom boxes indicate the third quartile and the first quartile, and the horizontal lines in the middle of the boxes indicate the median.

variance. *P* values <0.05 and partial $\eta^2 > 0.14$ (Richardson 2011) were required for the difference to be considered significant.

Results

Behavioral Findings

Figure 1A,B shows the grand-averaged voice F_0 contours in response to pitch perturbations of +200 and +500 cents across the c-TBS and sham conditions, illustrating c-TBS-induced effects on compensatory vocal responses to perturbed auditory feedback. A two-way RM-ANOVA revealed that c-TBS over left DLPFC led to significantly larger vocal responses as compared to the sham condition ($F(1, 20) = 32.284$, $P < 0.001$, partial $\eta^2 = 0.617$; see Fig. 1C). However, the magnitudes of vocal responses did not vary as a function of perturbation magnitude ($F(1, 20) = 2.377$, $P = 0.139$). Also, there was no significant interaction between stimulation condition and perturbation magnitude ($F(1, 20) = 0.066$, $P = 0.799$).

Regarding the latencies of vocal responses, there were no significant main effects of stimulation condition ($F(1, 20) = 2.986$, $P = 0.099$) and perturbation magnitude ($F(1, 20) = 0.133$, $P = 0.719$). A significant interaction, however, was found between stimulation condition and perturbation magnitude ($F(1, 20) = 6.374$, $P = 0.020$, partial $\eta^2 = 0.242$). Subsidiary analyses showed that the c-TBS condition led to significantly longer latencies of vocal compensations for +200 cents perturbations than the sham condition ($F(1, 20) = 7.501$, $P = 0.013$, partial $\eta^2 = 0.273$; see Fig. 1D), whereas this c-TBS effect was not significant for the +500 cents perturbations ($F(1, 20) = 0.030$, $P = 0.865$).

ERP Findings

Figure 2 shows the grand-averaged ERPs to pitch perturbations of +200 and +500 cents for the c-TBS and sham conditions across different electrode sites. As can be seen, the c-TBS condition elicited smaller P2 responses than the sham condition, whereas the N1 responses were not affected by c-TBS. This

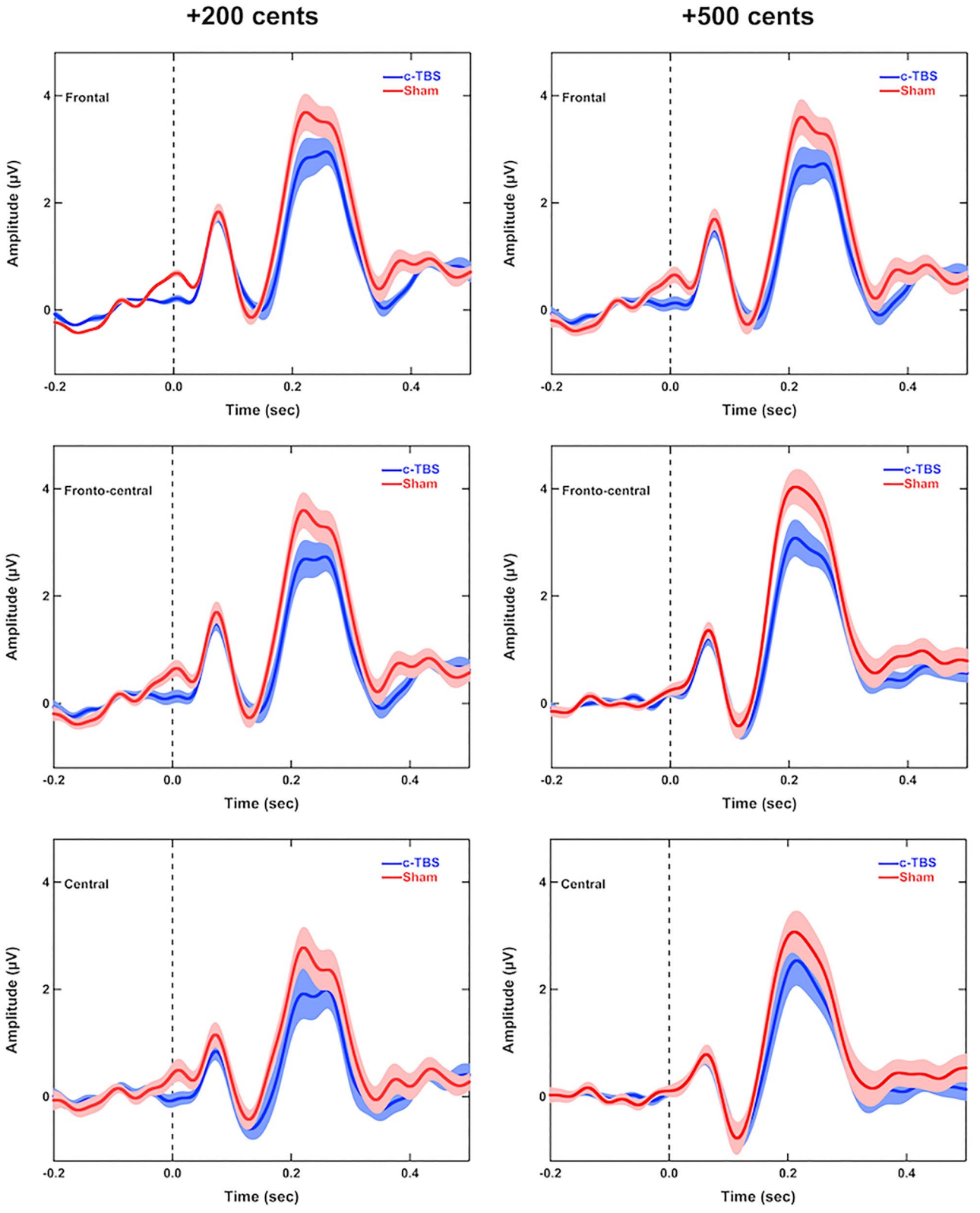

Figure 2. Grand-averaged ERPs to +200 (left panel) and +500 cents (right panel) perturbations for the c-TBS (blue solid lines) and sham (red solid lines) conditions in the frontal (top), fronto-central (middle), and central (bottom) regions. Highlighted areas indicate the standard errors of the mean ERPs. Vertical dashed lines indicate the onset of the pitch perturbation.

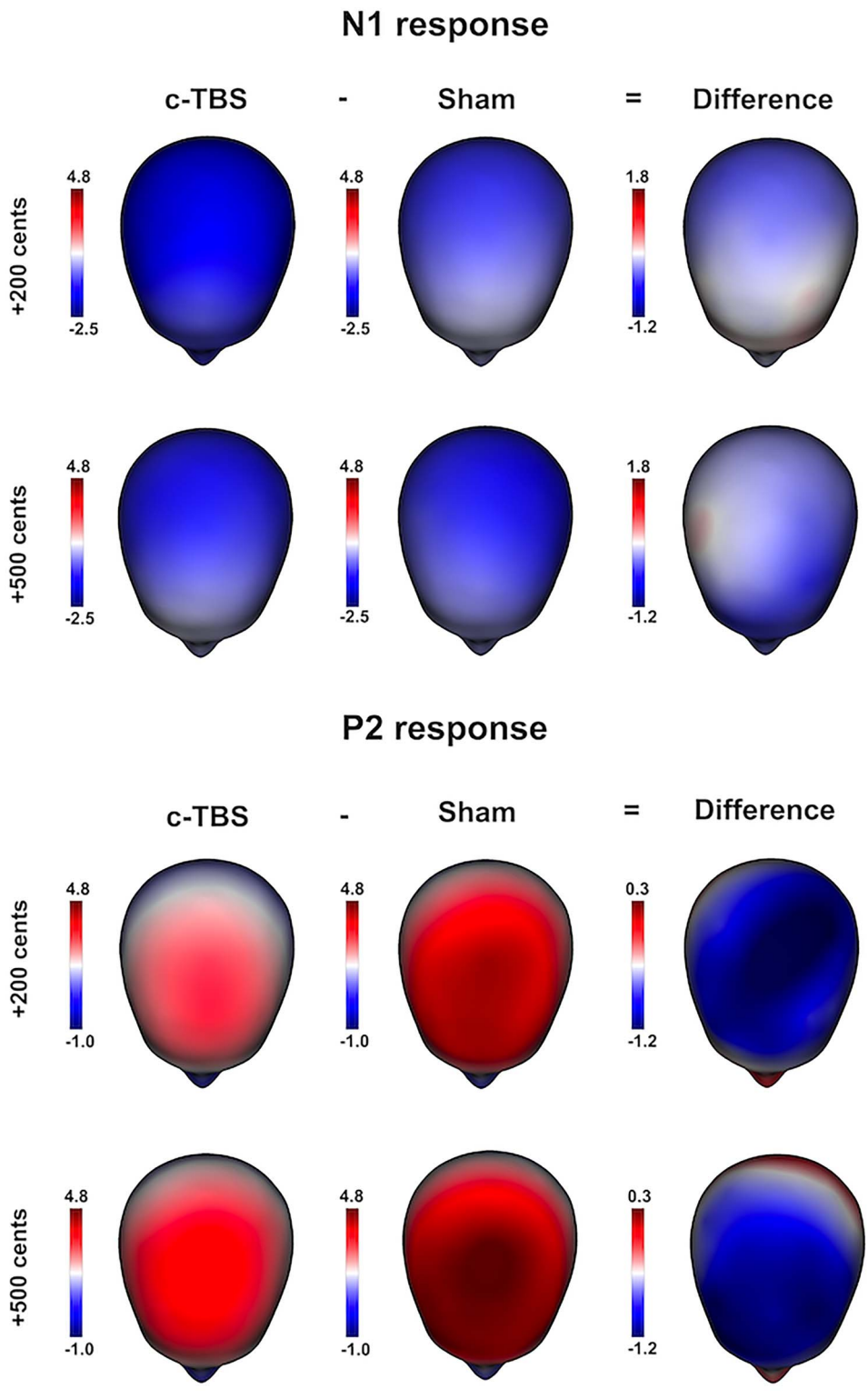

Figure 3. Topographical distribution maps of the N1 (top) and P2 (bottom) responses to +200 cents and +500 cents perturbations for the c-TBS and sham conditions as well as their differences.

c-TBS effect was further illustrated by the 3D topographical distributions of the N1 and P2 amplitudes shown in Figure 3; the c-TBS-induced effects were prominently pronounced in the P2 time window but subtle in the N1 time window.

A three-way RM-ANOVA conducted on the N1 amplitudes revealed a significant main effect of electrode site ($F(2, 40) = 9.030$, $P = 0.003$, partial $\eta^2 = 0.311$), which was primarily caused by significantly smaller (i.e., less negative) N1 amplitudes

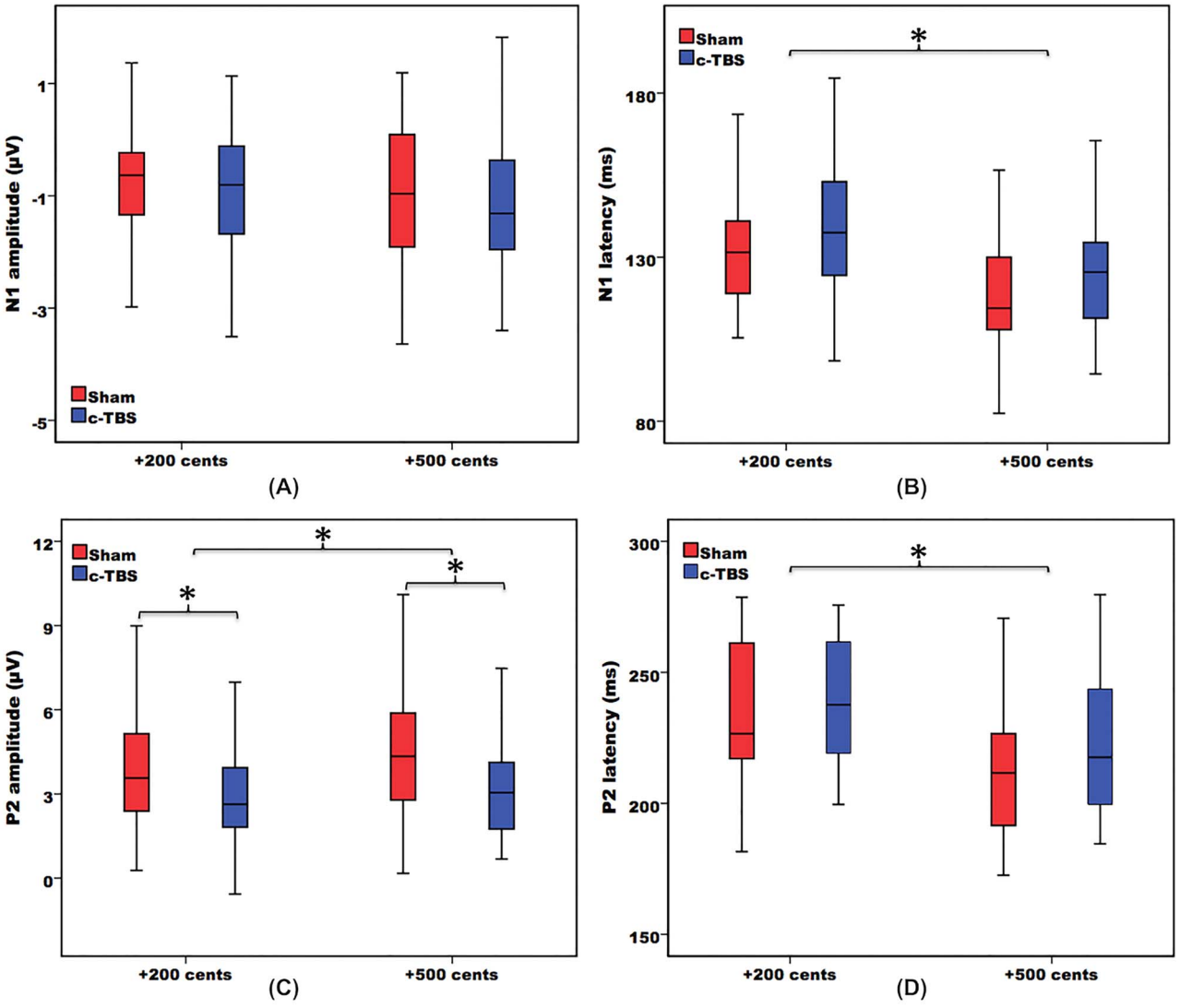

Figure 4. Box plots of the amplitudes and latencies of the N1 (*A* and *B*) and P2 (*C* and *D*) responses to +200 and +500 cents perturbations for the c-TBS and sham conditions. The asterisks indicate significant differences between the conditions. The top and bottom boxes indicate the third quartile and the first quartile, and the horizontal lines in the middle of the boxes indicate the median.

at the frontal electrodes relative to the fronto-central ($P < 0.001$) and central ($P = 0.036$) electrodes (see Fig. 4A). However, the N1 amplitudes did not vary as a function of stimulation condition ($F(1, 20) = 1.008$, $P = 0.327$) and perturbation magnitude ($F(1, 20) = 1.462$, $P = 0.241$). Interactions between any of three variables did not reach significance ($P > 0.15$).

For the N1 latencies, +500 cents perturbations elicited significantly faster N1 responses than +200 cents perturbations ($F(1, 20) = 20.590$, $P < 0.001$, partial $\eta^2 = 0.507$) (see Fig. 4B). There was a significant main effect of electrode site ($F(2, 40) = 5.384$, $P = 0.012$, partial $\eta^2 = 0.212$) as well as a significant interaction between stimulation condition and electrode site ($F(2, 40) = 3.763$, $P = 0.033$, partial $\eta^2 = 0.158$), whereas the main effect of stimulation condition was not significant ($F(1, 20) = 3.774$, $P = 0.066$). Subsidiary analyses revealed a significant main effect of electrode site for the c-TBS condition ($F(2, 40) = 8.458$, $P = 0.005$, partial $\eta^2 = 0.297$), where the N1 latencies at the central electrodes were significantly shorter than those at the frontal ($P = 0.020$) and fronto-central electrodes ($P = 0.009$). For the sham condition, however, the N1 latencies did not vary as a function of electrode site ($F(2, 40) = 2.924$, $P = 0.066$).

A three-way RM-ANOVA conducted on the P2 amplitudes revealed that c-TBS over the left DLPFC elicited significantly smaller P2 responses than the sham condition ($F(1, 20) = 9.255$, $P = 0.006$, partial $\eta^2 = 0.316$) and +500 cents perturbations elicited significantly larger P2 amplitudes than +200 cents perturbations ($F(1, 20) = 5.802$, $P = 0.026$, partial $\eta^2 = 0.225$; see Fig. 4C). There was also a significant main effect of electrode site ($F(2, 40) = 39.224$, $P < 0.001$, partial $\eta^2 = 0.662$), where P2 amplitudes at the central electrodes were significantly smaller as compared to the frontal ($P < 0.001$) and fronto-central electrodes ($P < 0.001$). Interactions between any of the three variables did not reach significance ($P > 0.2$).

Regarding the P2 latencies, significantly faster P2 responses were elicited by +500 cents perturbations relative to +200 cents perturbations ($F(1, 20) = 13.920$, $P = 0.001$, partial $\eta^2 = 0.410$;

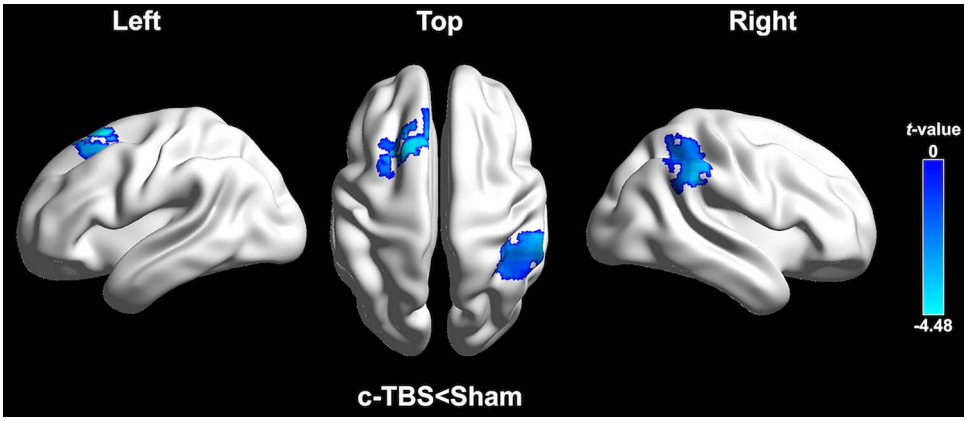

Figure 5. Grand-averaged sLORETA-based statistical nonparametric maps comparing the standardized current densities between c-TBS and sham conditions in the P2 time window. Results are projected onto lateral and top three-dimensional views of a standard anatomical template. Negative *t* values indicate decreased brain activity in response to pitch perturbations as a result of c-TBS over left DLPFC (corrected $P < 0.05$).

Table 1 sLORETA *t* statistics for maximum activations obtained from comparison between c-TBS and sham conditions in the P2 time window (MNI coordinates)

Condition	BA	Brain region	*t* value	X	Y	Z	P
c-TBS versus sham	8	Left SFG	−4.480	−20	15	65	0.0192
	40	Right IPL (SMG)	−4.251	55	−35	30	0.0214

Note: Displayed are *t* values for current density maxima; threshold for significance at corrected $P < 0.05$.

see Fig. 4D). The main effects of stimulation condition ($F(1, 20) = 2.413$, $P = 0.136$) and electrode site ($F(2, 40) = 3.246$, $P = 0.071$), however, were not significant. There were no significant interactions between any of the three variables ($P > 0.1$).

sLORETA Findings

sLORETA analyses were performed to examine the neural correlates of significant differences in the cortical P2 responses to pitch perturbations in voice auditory feedback between c-TBS and sham conditions. Considering that the effects of c-TBS on the P2 responses were not modulated by the size of pitch perturbations, the P2 responses to +200 and +500 cents perturbations were combined into a single dataset for the source estimation analyses. Differences in estimated current density between c-TBS and sham conditions in the N1 time window did not reach significance and were thus not illustrated. Figure 5 displays the sLORETA maps representing cortical regions that showed significant differences in brain activation between the c-TBS and sham conditions in the P2 time window. Table 1 shows the anatomical description and the MNI coordinates corresponding to these brain regions. The results revealed that significantly smaller P2 responses elicited by c-TBS relative to the sham condition were the result of reduced activity in the left superior frontal gyrus (SFG; BA 8, $P = 0.0192$) and right IPL (SMG; BA 40, $P = 0.0214$).

Discussion

The present study investigated the role of the left DLPFC in auditory feedback control of vocal production from a causal perspective. Consistent with our hypothesis, the behavioral results revealed that applying c-TBS over left DLPFC led to significantly enhanced vocal compensations for pitch perturbations relative to the sham condition. Modulatory effects of c-TBS on cortical activity were also observed, as reflected by reduced P2 responses that were source-localized in the left SFG and right IPL (SMG). These findings provide the first causal evidence for the role of the left DLPFC in auditory feedback control of vocal production. The top–down regulatory mechanisms mediated by the left DLPFC may exert inhibitory influences on auditory–vocal integration, through which vocal motor behaviors can be appropriately regulated without being excessively influenced by auditory feedback.

Comparison with Other Studies

Our results showed enhanced vocal compensations for pitch perturbations as a result of c-TBS over left DLPFC, indicating a role of the left DLPFC in the compensatory mechanisms that support vocal pitch regulation. This finding is in line with one recent study on patients with AD, where enhanced vocal compensations for pitch perturbations were correlated with reduced high-gamma activity in the left DLPFC (Ranasinghe et al. 2019). In addition, c-TBS-induced enhancement of vocal compensations was associated with reduced cortical P2 responses. This brain–behavior association was also found in other studies (Scheerer et al. 2013; Guo et al. 2017; Huang et al. 2019). Together with other studies that have shown that cortical P2 responses are modulated by attention control and working memory (Liu et al. 2015; Guo et al. 2017; Liu et al. 2018), the present study supports the notion that cortical P2 responses to pitch perturbations reflect higher-level cognitive aspects of auditory–vocal integration (Behroozmand et al. 2014; Guo et al. 2017).

Furthermore, source localization analysis revealed that the cortical P2 responses received contributions from the left SFG

and right IPL (SMG). Previous studies have also shown that, in parallel to contributions from the auditory cortex, additional current densities that originate from the IPL (SMG) contribute to the cortical P2 response to voice pitch perturbations (Huang et al. 2016; Guo et al. 2017; Li et al. 2019). This finding is also consistent with neuroimaging studies that associated activation of the IPL (SMG) with the generation of rapid vocal compensations for pitch perturbations (Toyomura et al. 2007; Kort et al. 2014; Kort et al. 2016; Behroozmand et al. 2017). Note that the IPL receives multisensory input and is believed to facilitate sensorimotor prediction and plasticity through its connections with prefrontal, primary motor, sensory, and somatosensory cortices as well as the cerebellum (Blakemore and Sirigu 2003; Rauschecker and Scott 2009). For speech production, the IPL has been implicated in speech motor learning, as evidenced by the involvement of the IPL (SMG) in sensorimotor adaptations to speech F1 perturbations (Shum et al. 2011). Together, these findings suggest that the IPL, particularly the SMG, is not only an important component in the network for speech motor learning but also plays a role in the online auditory–motor control of vocal production.

The magnitudes and latencies of vocal compensations did not vary as a function of perturbation magnitude. Across studies, the particular effect of perturbation magnitude on vocal responses has been quite varied. For example, one previous study found that larger pitch perturbations elicited larger vocal compensations during speech but not vowels in English speakers (Chen et al. 2007). In contrast, smaller vocal responses were found to be associated with larger pitch perturbations in Cantonese speakers but not in Mandarin speakers (Liu et al. 2010). As well, pitch perturbations of 350 cents or more elicited smaller vocal compensations than pitch perturbations of 250 cents or less in English speakers (Scheerer et al. 2013), while the magnitudes of vocal responses to 200 and 500 cents did not differ in Mandarin speakers (Guo et al. 2016; Li et al. 2018). The specific reason for the heterogeneous effects of perturbation magnitude on the vocal responses across the studies is not yet known, and they may be related to the specificity of task demands and the difference in language experience. At the cortical level, however, previous studies as well the present study have consistently shown that larger pitch perturbations elicit larger and faster N1 and P2 responses (Behroozmand et al. 2009; Liu et al. 2011; Scheerer et al. 2013). Note that overall the c-TBS effects on the neurobehavioral responses to pitch perturbations were not modulated by perturbation magnitude, but based on the above findings we cannot conclude that the mechanisms behind these effects are insensitive to the nature of the perturbations and more studies are needed to determine their relationship.

The Role of the DLPFC in Speech Motor Control

Current models of speech motor control have highlighted the importance of the inferior prefrontal cortex in sensorimotor integration for speech production. For example, the DIVA model argues that the feedback and feedforward control of speech production is initiated by a speech sound map stored in the left posterior IFG (Golfinopoulos et al. 2010). In the dual stream model, the left posterior IFG is a central part of a dorsal stream that is involved in translating acoustic speech signals into articulatory representations (Hickok and Poeppel 2007). Empirical evidence has shown activation of the IFG and its connectivity with temporal and parietal regions in the generation of vocal compensations for perceived perturbations in auditory feedback (Flagmeier et al. 2014; Behroozmand et al. 2015; Kort et al. 2016).

In contrast, less is known about the functional role of the DLPFC in speech motor control. The DLPFC is extensively connected to auditory and motor regions (Selemon and Goldman-Rakic 1988; Romanski et al. 1999) and has been implicated in both the perception and production of speech (Dehaene-Lambertz et al. 2002; Geranmayeh et al. 2012). Damage to the DLPFC leads to disorders in planning and programming of speech articulation (Ziegler 2002). In addition, patients with PD who received LSVT LOUD (Lee Silverman Voice Treatment), an intensive speech therapy program, for hypophonia exhibited increased activity in the DLPFC (Liotti et al. 2003) that was predictive of the treatment outcome (Narayana et al. 2010). These results suggest that LSVT LOUD may help patients with PD to increase their vocal loudness by altering activity in the DLPFC, which leads to changes in the motor and subcortical regions responsible for an effective speech outcome. Riecker et al. (2005) proposed a network that included the left DLPFC (including Broca's area), anterior insula, and right superior cerebellum that subserves motor preparation during syllable repetition. The role of the left DLPFC in that network may be to support higher-level aspects of speech motor control. In line with this idea, one recent study by Ranasinghe et al. (2019) showed that enhanced vocal compensation for pitch perturbations was predicted by reduced activity in the left DLPFC. The present study demonstrates a causal role of the DLPFC in speech motor control, as reflected by increased vocal compensations for pitch perturbations and reduced cortical activity in the left SFG and right IPL as a result of c-TBS over left DLPFC. Taken together, these findings provide strong evidence for the hypothesis that the DLPFC exerts top–down influences on the detection and correction of mismatches between the intended and actual vocal output.

The observed relationship between disrupted activity in the left DLPFC and enhanced vocal compensations for pitch perturbations reflects a top–down inhibitory effect on auditory feedback control of vocal production. One question then arises: Why does the left DLPFC inhibit vocal compensations for pitch perturbations? One important phenomenon that has been consistently observed is that perturbations in auditory feedback cannot be cancelled out completely; participants compensate for less than 20% of the size of the feedback perturbations (Burnett et al. 1998; Chen et al. 2007). Control theory suggests that fully compensating for feedback perturbations disrupts the stability of control systems (Franklin et al. 1994). Likewise, the degree of compensatory adjustments of vocal motor behaviors has to be controlled in an appropriately inhibitory manner in order to precisely and stably produce speech/vocal targets (Houde and Nagarajan 2011; Ranasinghe et al. 2017). Supportive evidence for this speculation comes from several studies that reported that singers were capable of completely ignoring pitch perturbations during singing (Zarate and Zatorre 2008; Zarate et al. 2010). On the other hand, the DLPFC serves as a critical source of inhibitory control (Garavan et al. 1999; Hampshire et al. 2010), an important function that inhibits reflex-like or inappropriate behavioral responses (Burle et al. 2004). There is evidence that has shown impaired inhibitory control processes due to decreased activity in the prefrontal and parietal regions (Barber et al. 2013). Note that, in the present study, c-TBS over left DLPFC led to reduced cortical P2 responses that source-localized in the left SFG and right IPL (SMG); both regions have been found to be involved in inhibitory control (Garavan et al. 1999). It is

thus plausible that the left DLPFC exerts top–down inhibitory influences on auditory feedback control of vocal production, allowing the audio–vocal system to partially compensate for feedback perturbations using an optimal strategy so that vocal targets can be produced with stability and precision. Disrupting cortical activity in the left DLPFC as a consequence of c-TBS, however, may impair this top–down regulatory mechanism and produce deficient inhibitory control over speech motor behaviors, which in turn results in enhanced vocal compensations for pitch perturbations.

Theoretical and Clinical Implications

Current models/theories of speech motor control describe the processes for monitoring and responding to auditory feedback during vocal production in an input-driven, bottom-up manner (Guenther et al. 2006; Golfinopoulos et al. 2010; Hickok et al. 2011; Houde and Nagarajan 2011; Guenther and Vladusich 2012). However, this control process has been shown to be modulated by higher-level cognitive functions such as attentional control and working memory (Tumber et al. 2014; Liu et al. 2015; Scheerer et al. 2016; Guo et al. 2017), reflecting top–down influences on auditory–vocal integration. According to the language comprehension model (Friederici 2012), bottom-up processing of information received by the auditory cortex occurs along the ventral pathway and reaches the frontal cortex. This bottom-up processing is done in parallel with top–down processing that originates in the frontal cortex and continues along the dorsal pathway to the temporal cortex. Similarly, the speech-motor-control network may also rely on the bottom-up, auditory-driven processes and top–down, prefrontally mediated processes. The present study provides significant insight into the neural mechanisms underlying the top–down modulation of vocal motor behaviors, and further studies should be conducted to unravel the neural circuits involved in these top–down regulatory mechanisms.

From a clinical perspective, the hypothesis that top–down mechanisms, mediated by the DLPFC, support vocal motor control accounts well for the relationship between impaired auditory–vocal integration and abnormal activity in the DLPFC in patients with AD (Ranasinghe et al. 2019) and PD (Pinto et al. 2004). The relationship between improved speech volume and increased activity in the DLPFC as a result of LSVT LOUD for patients with PD (Liotti et al. 2003; Narayana et al. 2010) suggests that modifying activity in the DLPFC may augment auditory feedback control of vocal production through its connections with the temporal and parietal regions. Repetitive TMS to the left DLPFC has been shown to improve language performance in patients with AD (Cotelli et al. 2011) and voice-related quality of life in patients with PD (Dias et al. 2006). Thus, modifying the activity in the DLPFC by noninvasive brain stimulation may potentially treat disorders in speech motor control by enhancing auditory–motor integration, which needs to be examined in future.

Funding

National Natural Science Foundation of China (Nos. 31371135, 81772439, 81972147); Guangdong Province Science and Technology Planning Project (No. 2017A050501014); Natural Science Foundation of Guangdong Province (No. 2018A030310037); Guangzhou Science and Technology Programme (No. 20160402 0115).

Note

Conflict of Interest: None declared.

References

Barber AD, Caffo BS, Pekar JJ, Mostofsky SH. 2013. Effects of working memory demand on neural mechanisms of motor response selection and control. *J Cogn Neurosci*. 25:1235–1248.

Bauer JJ, Mittal J, Larson CR, Hain TC. 2006. Vocal responses to unanticipated perturbations in voice loudness feedback: an automatic mechanism for stabilizing voice amplitude. *J Acoust Soc Am*. 119:2363–2371.

Behroozmand R, Ibrahim N, Korzyukov O, Robin DA, Larson CR. 2014. Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain Cogn*. 84:97–108.

Behroozmand R, Karvelis L, Liu H, Larson CR. 2009. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clin Neurophysiol*. 120:1303–1312.

Behroozmand R, Korzyukov O, Sattler L, Larson CR. 2012. Opposing and following vocal responses to pitch-shifted auditory feedback: evidence for different mechanisms of voice pitch control. *J Acoust Soc Am*. 132:2468–2477.

Behroozmand R, Phillip L, Johari K, Bonilha L, Rorden C, Hickok G, Fridriksson J. 2017. Sensorimotor impairment of speech auditory feedback processing in aphasia. *NeuroImage*. 165:102–111.

Behroozmand R, Shebek R, Hansen DR, Oya H, Robin DA, Howard MA 3rd, Greenlee JD. 2015. Sensory-motor networks involved in speech production and motor control: an fMRI study. *NeuroImage*. 109:418–428.

Blakemore SJ, Sirigu A. 2003. Action prediction in the cerebellum and in the parietal lobe. *Exp Brain Res*. 153:239–245.

Boersma P. 2001. Praat, a system for doing phonetics by computer. *Glot Int*. 5:341–345.

Brosnan MB, Wiegand I. 2017. The dorsolateral prefrontal cortex, a dynamic cortical area to enhance top-down attentional control. *J Neurosci*. 37:3445–3446.

Burle B, Vidal F, Tandonnet C, Hasbroucq T. 2004. Physiological evidence for response inhibition in choice reaction time tasks. *Brain Cogn*. 56:153–164.

Burnett TA, Freedland MB, Larson CR, Hain TC. 1998. Voice F0 responses to manipulations in pitch feedback. *J Acoust Soc Am*. 103:3153–3161.

Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF. 2013. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc Natl Acad Sci USA*. 110:2653–2658.

Chen SH, Liu H, Xu Y, Larson CR. 2007. Voice F0 responses to pitch-shifted voice feedback during English speech. *J Acoust Soc Am*. 121:1157–1163.

Cotelli M, Calabria M, Manenti R, Rosini S, Zanetti O, Cappa SF, Miniussi C. 2011. Improved language performance in Alzheimer disease following brain stimulation. *J Neurol Neurosurg Psychiatry*. 82:794–797.

Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L. 2002. Functional neuroimaging of speech perception in infants. *Science*. 298:2013–2015.

Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 134:9–21.

Devlin JT, Watkins KE. 2007. Stimulating language: insights from TMS. *Brain*. 130:610–622.

Dias AE, Barbosa ER, Coracini K, Maia F, Marcolm MA, Fregni F. 2006. Effects of repetitive transcranial magnetic stimulation on voice and speech in Parkinson's disease. *Acta Neurol Scand*. 113:92–99.

Edin F, Klingberg T, Johansson P, McNab F, Tegner J, Compte A. 2009. Mechanism for top-down control of working memory capacity. *Proc Natl Acad Sci USA*. 106:6802–6807.

Ferree TC, Luu P, Russell GS, Tucker DM. 2001. Scalp electrode impedance, infection risk, and EEG data quality. *Clin Neurophysiol*. 112:536–544.

Finkel S, Veit R, Lotze M, Friberg A, Vuust P, Soekadar S, Birbaumer N, Kleber B. 2019. Intermittent theta burst stimulation over right somatosensory larynx cortex enhances vocal pitch-regulation in nonsingers. *Hum Brain Mapp*. 40:2174–2187.

Flagmeier SG, Ray KL, Parkinson AL, Li K, Vargas R, Price LR, Laird AR, Larson CR, Robin DA. 2014. The neural changes in connectivity of the voice network during voice pitch perturbation. *Brain Lang*. 132:7–13.

Franklin GF, Powell JD, Emami-Naeini A. 1994. *Feedback Control of Dynamic Systems*. 3rd ed. Reading, MA: Addison-Wesley.

Friederici AD. 2012. The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn Sci*. 16:262–268.

Fuchs M, Kastner J, Wagner M, Hawes S, Ebersole JS. 2002. A standardized boundary element method volume conductor model. *Clin Neurophysiol*. 113:702–712.

Garavan H, Ross TJ, Stein EA. 1999. Right hemispheric dominance of inhibitory control: an event-related functional MRI study. *Proc Natl Acad Sci USA*. 96:8301–8306.

Geranmayeh F, Brownsett SL, Leech R, Beckmann CF, Woodhead Z, Wise RJ. 2012. The contribution of the inferior parietal cortex to spoken language production. *Brain Lang*. 121:47–57.

Golfinopoulos E, Tourville JA, Guenther FH. 2010. The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage*. 52:862–874.

Grossheinrich N, Rau A, Pogarell O, Hennig-Fast K, Reinl M, Karch S, Dieler A, Leicht G, Mulert C, Sterr A, et al. 2009. Theta burst stimulation of the prefrontal cortex: safety and impact on cognition, mood, and resting electroencephalogram. *Biol Psychiatry*. 65:778–784.

Guenther FH, Ghosh SS, Tourville JA. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang*. 96:280–301.

Guenther FH, Vladusich T. 2012. A neural theory of speech acquisition and production. *J Neurolinguistics*. 25:408–422.

Guo Z, Huang X, Wang M, Jones JA, Dai Z, Li W, Liu P, Liu H. 2016. Regional homogeneity of intrinsic brain activity correlates with auditory-motor processing of vocal pitch errors. *NeuroImage*. 142:565–575.

Guo Z, Wu X, Li W, Jones JA, Yan N, Sheft S, Liu P, Liu H. 2017. Top-down modulation of auditory-motor integration during speech production: the role of working memory. *J Neurosci*. 37:10323–10333.

Hampshire A, Chamberlain SR, Monti MM, Duncan J, Owen AM. 2010. The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage*. 50:1313–1319.

Herwig U, Satrapi P, Schonfeldt-Lecuona C. 2003. Using the international 10-20 EEG system for positioning of transcranial magnetic stimulation. *Brain Topogr*. 16:95–99.

Hickok G, Houde JF, Rong F. 2011. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*. 69:407–422.

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci*. 8:393–402.

Hoogendam JM, Ramakers GM, Di Lazzaro V. 2010. Physiology of repetitive transcranial magnetic stimulation of the human brain. *Brain Stimul*. 3:95–118.

Houde JF, Jordan MI. 1998. Sensorimotor adaptation in speech production. *Science*. 279:1213–1216.

Houde JF, Nagarajan SS. 2011. Speech production as state feedback control. *Front Hum Neurosci*. 5:82.

Hu H, Liu Y, Guo Z, Li W, Liu P, Chen S, Liu H. 2015. Attention modulates cortical processing of pitch feedback errors in voice control. *Sci Rep*. 5:7812.

Huang X, Chen X, Yan N, Jones JA, Wang EQ, Chen L, Guo Z, Li W, Liu P, Liu H. 2016. The impact of Parkinson's disease on the cortical mechanisms that support auditory-motor integration for voice control. *Hum Brain Mapp*. 37: 4248–4261.

Huang X, Fan H, Li J, Jones JA, Wang EQ, Chen L, Chen X, Liu H. 2019. External cueing facilitates auditory-motor integration for speech control in individuals with Parkinson's disease. *Neurobiol Aging*. 76:96–105.

Huang YZ, Edwards MJ, Rounis E, Bhatia KP, Rothwell JC. 2005. Theta burst stimulation of the human motor cortex. *Neuron*. 45:201–206.

Knight RT, Scabini D, Woods DL. 1989. Prefrontal cortex gating of auditory transmission in humans. *Brain Res*. 504: 338–342.

Kort NS, Cuesta P, Houde JF, Nagarajan SS. 2016. Bihemispheric network dynamics coordinating vocal feedback control. *Hum Brain Mapp*. 37:1474–1485.

Kort NS, Nagarajan SS, Houde JF. 2014. A bilateral cortical network responds to pitch perturbations in speech feedback. *NeuroImage*. 86:525–535.

Korzyukov O, Karvelis L, Behroozmand R, Larson CR. 2012a. ERP correlates of auditory processing during automatic correction of unexpected perturbations in voice auditory feedback. *Int J Psychophysiol*. 83:71–78.

Korzyukov O, Sattler L, Behroozmand R, Larson CR. 2012b. Neuronal mechanisms of voice control are affected by implicit expectancy of externally triggered perturbations in auditory feedback. *PLoS One*. 7:e41216.

Lega C, Stephan MA, Zatorre RJ, Penhune V. 2016. Testing the role of dorsal premotor cortex in auditory-motor association learning using transcranical magnetic stimulation (TMS). *PLoS One*. 11:e0163380.

Li J, Hu H, Chen N, Jones JA, Wu D, Liu P, Liu H. 2018. Aging and sex influence cortical auditory-motor integration for speech control. *Front Neurosci*. 12:749.

Li W, Chen Z, Liu P, Zhang B, Huang D, Liu H. 2013. Neurophysiological evidence of differential mechanisms involved in producing opposing and following responses to altered auditory feedback. *Clin Neurophysiol*. 124:2161–2171.

Li W, Zhuang J, Guo Z, Jones JA, Xu Z, Liu H. 2019. Cerebellar contribution to auditory feedback control of speech production: evidence from patients with spinocerebellar ataxia. *Hum Brain Mapp*. 40:4748–4758.

Liotti M, Ramig LO, Vogel D, New P, Cook CI, Ingham RJ, Ingham JC, Fox PT. 2003. Hypophonia in Parkinson's disease: neural correlates of voice treatment revealed by PET. *Neurology*. 60:432–440.

Liu H, Larson CR. 2007. Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *J Acoust Soc Am*. 122:3671–3677.

Liu H, Meshman M, Behroozmand R, Larson CR. 2011. Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clin Neurophysiol*. 122:951–957.

Liu H, Wang EQ, Chen Z, Liu P, Larson CR, Huang D. 2010. Effect of tonal native language on voice fundamental frequency responses to pitch feedback perturbations during vocalization. *J Acoust Soc Am*. 128:3739–3746.

Liu Y, Fan H, Li J, Jones JA, Liu P, Zhang B, Liu H. 2018. Auditory-motor control of vocal production during divided attention: behavioral and ERP correlates. *Front Neurosci*. 12:113.

Liu Y, Hu H, Jones JA, Guo Z, Li W, Chen X, Liu P, Liu H. 2015. Selective and divided attention modulates auditory-vocal integration in the processing of pitch feedback errors. *Eur J Neurosci*. 42:1895–1904.

Mansouri FA, Tanaka K, Buckley MJ. 2009. Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nat Rev Neurosci*. 10:141–152.

Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Woods R, Paus T, Simpson G, Pike B, et al. 2001. A probabilistic atlas and reference system for the human brain: international consortium for brain mapping (ICBM). *Philos Trans R Soc Lond Ser B Biol Sci*. 356:1293–1322.

Mitchell TV, Morey RA, Inan S, Belger A. 2005. Functional magnetic resonance imaging measure of automatic and controlled auditory processing. *Neuroreport*. 16:457–461.

Mulert C, Jager L, Schmitt R, Bussfeld P, Pogarell O, Moller HJ, Juckel G, Hegerl U. 2004. Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *NeuroImage*. 22:83–94.

Murakami T, Ugawa Y, Ziemann U. 2013. Utility of TMS to understand the neurobiology of speech. *Front Psychol*. 4:446.

Murdoch BE, Barwood CH. 2013. Non-invasive brain stimulation: a new frontier in the treatment of neurogenic speech-language disorders. *Int J Speech Lang Pathol*. 15:234–244.

Narayana S, Fox PT, Zhang W, Franklin C, Robin DA, Vogel D, Ramig LO. 2010. Neural correlates of efficacy of voice therapy in Parkinson's disease identified by performance-correlation analysis. *Hum Brain Mapp*. 31:222–236.

Natke U, Donath TM, Kalveram KT. 2003. Control of voice fundamental frequency in speaking versus singing. *J Acoust Soc Am*. 113:1587–1593.

Parkinson AL, Flagmeier SG, Manes JL, Larson CR, Rogers B, Robin DA. 2012. Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage*. 61:314–322.

Parrell B, Agnew Z, Nagarajan S, Houde J, Ivry RB. 2017. Impaired feedforward control and enhanced feedback control of speech in patients with cerebellar degeneration. *J Neurosci*. 37:9249–9258.

Patel SP, Kim JH, Larson CR, Losh M. 2019. Mechanisms of voice control related to prosody in autism spectrum disorder and first-degree relatives. *Autism Res*. 12:1192–1210.

Pinto S, Thobois S, Costes N, Le Bars D, Benabid AL, Broussolle E, Pollak P, Gentil M. 2004. Subthalamic nucleus stimulation and dysarthria in Parkinson's disease: a PET study. *Brain*. 127:602–615.

Ranasinghe KG, Gill JS, Kothare H, Beagle AJ, Mizuiri D, Honma SM, Gorno-Tempini ML, Miller BL, Vossel KA, Nagarajan SS, et al. 2017. Abnormal vocal behavior predicts executive and memory deficits in Alzheimer's disease. *Neurobiol Aging*. 52:71–80.

Ranasinghe KG, Kothare H, Kort N, Hinkley LB, Beagle AJ, Mizuiri D, Honma SM, Lee R, Miller BL, Gorno-Tempini ML, et al. 2019. Neural correlates of abnormal auditory feedback processing during speech production in Alzheimer's disease. *Sci Rep*. 9:5686.

Rauschecker JP, Scott SK. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci*. 12:718–724.

Restle J, Murakami T, Ziemann U. 2012. Facilitation of speech repetition accuracy by theta burst stimulation of the left posterior inferior frontal gyrus. *Neuropsychologia*. 50:2026–2031.

Richardson JTE. 2011. Eta squared and partial eta squared as measures of effect size in educational research. *Educ Res Rev*. 6:135–147.

Riecker A, Mathiak K, Wildgruber D, Erb M, Hertrich I, Grodd W, Ackermann H. 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology*. 64:700–706.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci*. 2:1131–1136.

Scheerer NE, Behich J, Liu H, Jones JA. 2013. ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience*. 240:176–185.

Scheerer NE, Tumber AK, Jones JA. 2016. Attentional demands modulate sensorimotor learning induced by persistent exposure to changes in auditory feedback. *J Neurophysiol*. 115:826–832.

Selemon LD, Goldman-Rakic PS. 1988. Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network subserving spatially guided behavior. *J Neurosci*. 8:4049–4068.

Shum M, Shiller DM, Baum SR, Gracco VL. 2011. Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *Eur J Neurosci*. 34:1817–1822.

Teng S, Guo Z, Peng H, Xing G, Chen H, He B, McClure MA, Mu Q. 2017. High-frequency repetitive transcranial magnetic stimulation over the left DLPFC for major depression: session-dependent efficacy: a meta-analysis. *Eur Psychiatry*. 41: 75–84.

Toyomura A, Koyama S, Miyamaoto T, Terao A, Omori T, Murohashi H, Kuriki S. 2007. Neural correlates of auditory feedback control in human. *Neuroscience*. 146:499–503.

Tumber AK, Scheerer NE, Jones JA. 2014. Attentional demands influence vocal compensations to pitch errors heard in auditory feedback. *PLoS One*. 9:e109968.

Xia M, Wang J, He Y. 2013. BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One*. 8:e68910.

Zarate JM, Wood S, Zatorre RJ. 2010. Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia*. 48:607–618.

Zarate JM, Zatorre RJ. 2008. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage*. 40:1871–1887.

Ziegler W. 2002. Psycholinguistic and motor theories of apraxia of speech. *Semin Speech Lang*. 23:231–244.

Zumsteg D, Lozano AM, Wennberg RA. 2006. Depth electrode recorded cerebral responses with deep brain stimulation of the anterior thalamus for epilepsy. *Clin Neurophysiol*. 117:1602–1609.