



# Decreased Gray-Matter Volume in Insular Cortex as a Correlate of Singers' Enhanced Sensorimotor Control of Vocal Production

Wenda Wang<sup>1,2†</sup>, Lirao Wei<sup>3†</sup>, Na Chen<sup>4†</sup>, Jeffery A. Jones<sup>5</sup>, Gaolang Gong<sup>6\*</sup> and Hanjun Liu<sup>1,7\*</sup>

<sup>1</sup> Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, <sup>2</sup> Department of Rehabilitation, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China, <sup>3</sup> Department of Music, Guangdong University of Education, Guangzhou, China, <sup>4</sup> Department of Rehabilitation, Zhujiang Hospital, Southern Medical University, Guangzhou, China, <sup>5</sup> Psychology Department and Laurier Centre for Cognitive Neuroscience, Wilfrid Laurier University, Waterloo, ON, Canada, <sup>6</sup> State Key Laboratory of Cognitive Neuroscience and Learning and IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China, <sup>7</sup> Guangdong Provincial Key Laboratory of Brain Function and Disease, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China

## OPEN ACCESS

### Edited by:

Yi Du,  
Institute of Psychology (CAS), China

### Reviewed by:

Boris Kleber,  
Aarhus University, Denmark  
Xing Tian,  
New York University Shanghai, China

### \*Correspondence:

Gaolang Gong  
gaolang.gong@bnu.edu.cn  
Hanjun Liu  
lhajun@mail.sysu.edu.cn

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Auditory Cognitive Neuroscience,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 27 January 2019

**Accepted:** 22 July 2019

**Published:** 02 August 2019

### Citation:

Wang W, Wei L, Chen N,  
Jones JA, Gong G and Liu H (2019)  
Decreased Gray-Matter Volume  
in Insular Cortex as a Correlate  
of Singers' Enhanced Sensorimotor  
Control of Vocal Production.  
*Front. Neurosci.* 13:815.  
doi: 10.3389/fnins.2019.00815

Accumulating evidence has shown enhanced sensorimotor control of vocal production as a consequence of extensive singing experience. The neural basis of this ability, however, is poorly understood. Given that the insula mediates motor aspects of vocal production, the present study investigated structural plasticity in insula induced by singing experience and its link to auditory feedback control of vocal production. Voxel-based morphometry (VBM) was used to examine the differences in gray matter (GM) volume in the insula of 21 singers and 21 non-singers. An auditory feedback perturbation paradigm was used to examine the differences in auditory-motor control of vocal production between singers and non-singers. Both groups vocalized sustained vowels while hearing their voice unexpectedly pitch-shifted  $-50$  or  $-200$  cents (200 ms duration). VBM analyses showed that singers exhibited significantly lower GM volumes in the bilateral insula than non-singers. When exposed to pitch perturbations in voice auditory feedback, singers involuntarily compensated for pitch perturbations in voice auditory feedback to a significantly lesser degree than non-singers. Moreover, across the two sizes of pitch perturbations, the magnitudes of vocal compensations were positively correlated with the total regional GM volumes in the bilateral insula. These results indicate that extensive singing training leads to decreased GM volumes in insula and suggest that morphometric plasticity in insula contributes to the enhanced sensorimotor control of vocal production observed in singers.

**Keywords:** auditory feedback, speech motor control, voxel-based morphology, insula, singing

## INTRODUCTION

The goal of speech motor control is intelligible speech sounds. These sounds, perceived as auditory feedback, provide critical information that allows the brain to detect errors in vocal output and initiate motor commands that correct for them (Guenther, 2006; Houde and Chang, 2015). Speakers have been consistently shown to compensate for perturbations heard in their voice

fundamental frequency ( $F_0$ ), intensity, and formant frequency (Burnett et al., 1998; Houde and Jordan, 1998; Jones and Munhall, 2005; Bauer et al., 2006; Purcell and Munhall, 2006; Liu and Larson, 2007). This compensatory process engages a large complex neural network of sensory, motor, and cognitive systems (Tourville et al., 2008; Parkinson et al., 2012; Chang et al., 2013; Kort et al., 2014; Behroozmand et al., 2015, 2017). Our understanding of sensorimotor integration for voice control, however, is still far from clear.

A growing body of literature has shown that, across participants, the ability to compensate for perturbations in voice auditory feedback is not equal, but rather varies as a function of expertise like singing. For example, a series of singing studies found that, when auditory feedback was perturbed in pitch during syllable singing, singers suppressed their vocal compensations for pitch perturbations to a lesser degree and were even able to completely ignore large pitch perturbations (i.e., produced almost no vocal compensation) compared to non-singers (Zarate and Zatorre, 2005, 2008; Zarate et al., 2010). This behavioral difference between the two groups was accompanied by distinct neural networks: singers showed increased activity in the inferior parietal lobule (IPL), superior temporal gyrus (STG), superior temporal sulcus (STS), and insula while non-singers recruited the anterior cingulate cortex (ACC), premotor cortex (PMC), and supramarginal gyrus (SMG) (Zarate and Zatorre, 2005, 2008). Music experience can also influence the reaction time that participants correct for voice pitch feedback errors, as reflected by longer reaction time for highly skilled singers than for moderately skilled singers (Grell et al., 2009). In one sensorimotor adaptation study by Jones and Keough (2008), singers not only produced smaller vocal compensations than non-singers during the pitch perturbation phase but also exhibited a larger aftereffect as reflected by higher  $F_0$  values when auditory feedback returned normal compared to the baseline phase (i.e., no perturbation). Similarly, Kleber et al. (2017) found singers' pitch matching accuracy to be significantly more preserved than non-singers when their auditory feedback was masked by noise. These results suggest that singing experience may lead to decreased reliance on auditory feedback and that instead, trained singers may rely relatively more on feedforward control mechanisms or the "acquired neuromuscular memory of pitch" (Murbe et al., 2004) to produce the vocal targets with increased precision.

Note that the aforementioned studies of speech motor control instructed musically trained participants to sing the syllables, whereas other studies that involved non-musicians typically employed speaking tasks (i.e., vocalizing the vowel sounds) (Burnett et al., 1998; Liu and Larson, 2007; Liu et al., 2011; Scheerer et al., 2013; Kort et al., 2014; Behroozmand et al., 2015). Despite the fundamental similarity in the way pitch is used for singing and speaking, singing requires more accurate encoding of pitch information and a higher level of vocal motor control than does speaking (Zatorre and Baum, 2012), which leads to the additional recruitment of right-hemisphere brain regions that include the STG and inferior frontal gyrus (IFG) as well as the insula (Riecker et al., 2000; Ozdemir et al., 2006). One behavioral study showed that singing the syllables leads to

larger vocal compensations for pitch perturbations than speaking the syllables in non-musicians (Natke et al., 2003). It should be noted that the previously reported vocal responses produced by singers were measured 1,900–3,000 ms after the onset of the 3-s-long pitch perturbation (Zarate and Zatorre, 2008; Zarate et al., 2010). It has been suggested that these relatively late responses are voluntary and reflect a conscious strategy to oppose perceived changes in voice auditory feedback (Burnett et al., 1998; Hain et al., 2000). In contrast, other studies that involved non-musicians measured the vocal responses 50–400 ms after the onset of a pitch perturbation that was 200–400 ms long (Chen et al., 2007; Liu and Larson, 2007; Parkinson et al., 2012; Scheerer et al., 2013; Kort et al., 2014). These early responses are reflex-like or involuntary and are unlikely to be consciously modified. Therefore, singers' ability to suppress compensatory adjustments of their vocal motor behaviors may be influenced by the specificity of the task demands (i.e., singing vs. speaking) and the nature of the vocal responses (i.e., voluntary vs. involuntary). This idea is supported by one study by Behroozmand et al. (2014) who showed that musicians with absolute pitch (AP) and relative pitch (RP) did not suppress their vocal responses, but instead compensated for the pitch perturbations to a relatively larger degree than non-musicians when they vocalized the vowel sounds. However, because Behroozmand et al. (2014) included both singers and instrumentalists in their AP and RP musician groups, it remains unclear how the integration of auditory feedback into ongoing voice control is modulated by singing experience.

Despite considerable research on experience-dependent functional changes in the cortical representations of sensorimotor integration for speech (Riecker et al., 2005; Zarate and Zatorre, 2005, 2008; Ozdemir et al., 2006; Kleber et al., 2010; Zarate et al., 2010), structural plasticity as a function of singing expertise and the assessment of its relationship with speech motor control have rarely been investigated. Numerous studies have shown that becoming a proficient instrumentalist leads to increased gray matter (GM) volume and cortical thickness in the auditory- and motor-related regions as well as reorganization of white matter (WM) (Schneider et al., 2002; Gaser and Schlaug, 2003; Bengtsson et al., 2005; Bermudez et al., 2009; Steele et al., 2013; Groussard et al., 2014). In contrast, only a few studies have investigated the structural plasticity induced by singing experience. In one diffusion tensor imaging (DTI) study by Halwani et al. (2011), both trained singers and instrumentalists exhibited larger WM tract volumes than non-musicians in the right arcuate fasciculus (AF), which connects fronto-temporal, sensorimotor, and inferior parietal regions (Glasser and Rilling, 2008). In a more recent voxel-based morphometry (VBM) study by Kleber et al. (2016), singers exhibited larger GM volumes in the right primary and secondary somatosensory cortices (S1 and S2), rostral SMG, and primary auditory cortex (A1), regions that have been shown to be active in compensating for perturbed voice  $F_0$  during speaking (Tourville et al., 2008; Chang et al., 2013; Behroozmand et al., 2015; Kort et al., 2016). Therefore, it is reasonable to hypothesize that long-term auditory-vocal training leads to structural changes in brain regions that are functionally relevant for sensorimotor control of

speech production, and that the assessment of their relationship may provide a window into the structural basis of speech motor control as a function of singing expertise.

The present VBM study investigated the neuroanatomical correlates of auditory feedback control of vocal pitch regulation in singers with the intention of (1) revealing whether GM volume differences in an *a priori* region of interest (ROI) would exist between singers and non-singers, and (2) examining the relationship between GM volume in the selected ROI and participants' vocal compensations for pitch perturbations in auditory feedback. The ROI selected for the present study was the insular cortex bilaterally. The insula is a complex structure that has a wide array of cortical connections with the frontal (e.g., PMC), temporal (e.g., STG, STS), and parietal (e.g., IPL) regions (Shelley and Trimble, 2004; Ghaziri et al., 2017). These widespread connections between the insula and other brain regions provide a role for the insula in a variety of sensorimotor integration and cognitive functions, including speech/language processing (Dronkers, 1996), central audition (Bamiou et al., 2006), body awareness (Craig, 2009), salience detection (Critchley et al., 2004), and affective processes (Deen et al., 2011; Uddin et al., 2014). More specifically, evidence from lesion and neuroimaging studies has demonstrated the involvement of the insula in the motor control of speech production. For example, lesions to the insula can lead to deficits in speech articulation and motor planning, such as apraxia of speech, reduced fluency, and impairments with articulatory movement (Dronkers, 1996; Bates et al., 2003; Ackermann and Riecker, 2010). Neuroimaging studies of healthy populations have identified increased activation in the insula during the production of compensatory vocal responses to pitch perturbations during singing (Zarate and Zatorre, 2008; Zarate et al., 2010) and speaking (Toyomura et al., 2007; Behroozmand et al., 2015; Kort et al., 2016). Moreover, Kleber et al. (2013) reported that, when asked to maintain pitch-matching performance after anesthesia of the vocal tracts, singers exhibited decreased activity in the right anterior insula and decreased connectivity between the insula and the auditory and somatosensory regions. In a subsequent study by Kleber et al. (2017), singers exhibited increased activity in the right anterior insula and increased connectivity between the insula and the SMG when they sang in the absence of auditory feedback. These findings suggest that the insula serves as a critical hub for the coordination of large-scale brain networks involved in integrating sensory, somatosensory, and motor information for speech motor control (Zarate and Zatorre, 2008).

Based on previous research that has shown structural neuroplasticity as a function of singing experience and the contribution of the insula to speech motor control, we predicted that singers and non-singers would significantly differ in GM volume in the selected ROI (i.e., bilateral insular cortex) and auditory feedback control of vocal pitch production. Furthermore, we predicted that a significant correlation would exist between GM volume in the insula and participants' ability to compensate for vocal pitch perturbations. The results confirmed our hypotheses; we observed reduced regional GM volume in the

bilateral insula and decreased involuntary vocal compensations for pitch perturbations in singers and positive correlations between GM volumes in those regions and the magnitudes of vocal compensations.

## MATERIALS AND METHODS

### Subjects

A total of 42 college students participated in the experiment. A group of 21 female classically trained singers [19–29 years, mean = 24.09, standard deviation (*SD*) = 2.19], consisting of 4 undergraduate students and 17 graduate students, were recruited from the Department of Music at South China Normal University. The participants took their formal singing lessons from the average age of  $12 \pm 5$  years (range: 4–21 years) and studied for an average of  $12 \pm 5$  years (range: 5–20 years). Some of the trained singers also played the piano. In addition, all singers reported that they did not possess AP. Twenty-one female college students (18–27 years, mean = 23.17, *SD* = 2.65) without previous vocal training or instrumental playing experience were recruited from Sun Yat-sen University (4 undergraduate and 17 graduate students) and assigned to the non-singers group. The two groups were matched in age ( $t = 1.139$ ,  $p = 0.268$ ), gender, and education. Based on self-report measures of alcohol and tobacco use, all participants were classified as non-drinker (0 units of alcohol per week) and non-smokers (0 cigarettes per day). All participants were right-handed and native Mandarin speakers. They reported no history of speech, language, hearing, and neurological disorders. All participants passed a binaural hearing screening at the threshold of 25 dB HL for pure tone frequencies of 0.5–4 kHz. They received monetary compensation for their participation and gave written informed consent in compliance with a research protocol approved by the Institution Review Board of The First Affiliated Hospital at Sun Yat-sen University of China.

### Structural MRI Acquisition

Structural MRI data were acquired on a Siemens Magnetom 3T Trio Tim MRI scanner (Erlangen, Germany) located at South China Normal University. During acquisition, all participants were required to lie still and stay awake with their eyes closed. The high-resolution anatomical images were acquired using T1-weighted 3D Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence with the following parameters: TR (repetition time) = 2,300 ms; TE (echo time) = 3.24 ms; flip angle = 9°; FOV (field of view) =  $256 \times 256$  mm<sup>2</sup>; slices thickness = 1 mm; and voxel size =  $1 \times 1 \times 1$  mm.

### Vocal Data Acquisition

After the MRI data acquisition, both singers and non-singers participated in a vocal production experiment using the frequency-altered feedback (FAF) paradigm. They were instructed to vocalize the vowel sound /u/ for approximately 5–6 s and to speak at their comfortable pitch and loudness level. During each vocalization, participants' voices were shifted down in pitch by 50 or 200 cents (100 cents equals one semitone) five times.

The duration of each pitch perturbation was fixed at 200 ms. The two sizes of pitch perturbations were presented pseudorandomly across all participants; the initial pitch perturbation occurred with a delay of 500–1,000 ms relative to the vocal onset, and the succeeding stimuli were presented with an inter-stimulus interval of 700–900 ms. Each participant was required to take a break of 2–3 s between successive vocalizations and produced 40 consecutive vocalizations. A total of 200 trials were thus collected, including 100 trials for the –50 cents condition and 100 trials for the –200 cents condition.

Throughout the experiment, the vocal data were collected from participants while they sat in a sound-treated booth. The voice feedback was calibrated to be 10 dB SPL higher than that of participant's vocal output using a Zwisklocki coupler and a Brüel & Kjaer sound level meter (model 2250) to reduce the influence of the air-born and bone-conducted voice feedback (Larson et al., 2008; Patel et al., 2014). During the experiment, the voice signals were recorded via a dynamic microphone (DM2200, Takstar Inc.), amplified with a MOTU Ultralite Mk3 Firewire audio interface, and pitch-shifted by an Eventide Eclipse Harmonizer controlled by a Max/MSP software program (v.5.0 by Cycling 74). The pitch-shifted voice signals were then amplified by an ICON NeoAmp headphone amplifier and presented to participants through insert earphones (ER1, Etymotic Research Inc.). Transistor–transistor logic (TTL) pulses were generated by the Max/MSP software program to mark each pitch shift event for averaging the vocal trials. The acoustic data and the TTL pulses were recorded at a sampling frequency of 10 kHz by a PowerLab A/D converter (model ML880, AD Instruments) using LabChart software (v.7.0 by AD Instruments). Note that the scalp-recorded electroencephalography (EEG) data were also recorded using a 64-electrode Geodesic Sensor Net through a Net Amps 300 amplifier (Electrical Geodesics Inc., Eugene, OR, United States), but the ERP results are not reported here.

## Data Analysis

### Vocal Data Analysis

Compensatory vocal responses to pitch perturbations were measured using IGOR PRO software (v.6.0 by Wavemetrics Inc.) using the event-related averaging technique (Liu and Larson, 2007; Larson et al., 2008). The voice  $F_0$  contours were extracted using Praat software (Boersma, 2001) and converted to the cent scale with the following formula: cents =  $100 \times (12 \times \log_2(F_0/\text{reference}))$  [reference = 195.997 Hz (G3 note)]. The voice contours were then segmented into trials with a window of 200 ms before and 700 ms after the perturbation onset. A visual inspection was performed to remove those trials that were contaminated by errors in vocal production or signal processing. Those artifact-free trials that opposed the direction of pitch perturbations were averaged and baseline-corrected to generate an overall compensatory vocal response for each condition (Li et al., 2013). The magnitude and latency of a vocal response were defined as the  $F_0$  value in cents and time in ms when the voice  $F_0$  contours reached their maximum values, respectively. We chose a time window of 50–400 ms after the perturbation onset to detect the prominent peaks

of the involuntary vocal responses based on the present and previous studies (Burnett et al., 1998; Behroozmand et al., 2014; Tang et al., 2018).

### MRI Data Processing

Structural MRI image data were processed using the Statistic Parametric Mapping software (SPM 12<sup>1</sup>). The CAT12 toolbox<sup>2</sup> implemented in SPM 12 was used for VBM analysis (Ashburner and Friston, 2000). First, all T1-weighted images were segmented into GM, WM, and cerebral spinal fluid (CSF) for the calculation of the overall tissue volume (GM, WM, and CSF) and total intracranial volume (TIV) in the native space. An internal GM threshold of 0.2 was used to exclude those artifacts on the gray-white-matter border. All of the segmented tissues were then registered to the standard Montreal Neurological Institute template in SPM12 using the affine registration algorithm. The Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) toolbox was used to refine the inter-subject registration of all participants' GM and WM. A statistical quality check procedure was performed using the CAT12 toolbox to assess the homogeneity of the GM tissues when the preprocessing pipeline was completed. Finally, normalized GM tissue segments for each participant were smoothed with an 8-mm full width at half maximum (FWHM) Gaussian filter.

### Statistical Analysis

The magnitudes and latencies of vocal compensations for pitch perturbations were subjected to repeated-measures analyses of variance (RM-ANOVAs). Group (singers vs. non-singers) was chosen as a between-subject factor, while stimulus magnitude (–50 cents vs. –200 cents) was chosen as a within-subject factor. Significant higher-order interactions between two variables led to subsidiary RM-ANOVAs. Probability values were corrected for multiple degrees of freedom when violations of sphericity occurred. The effect size indexed by  $\eta_p^2$  was calculated to indicate the size of differences across the conditions. An alpha level of  $p < 0.05$  was considered to be significant.

Anatomical group differences with respect to GM volume were assessed in *ad hoc* defined ROIs, focusing on the left and right insular cortex. In this ROI analysis, the bilateral insular cortices were selected as masks from the automated anatomical labeling (AAL) template (Tzourio-Mazoyer et al., 2002) using the Data Processing and Analysis of Brain Imaging (DPABI) MATLAB toolbox implemented in SPM 12 (Yan et al., 2016). A voxel-based comparison with independent two-sample *t*-tests was performed to detect differences between singers and non-singers within the mask. The significance of group differences in the ROI was estimated to correct for multiple comparisons using the Gaussian random field (GRF) theory (voxel-level significance:  $p < 0.001$ ; cluster-level significance:  $p < 0.05$ ) with covariates of age and TIV included. We extracted the total GM volumes of each significant cluster within each ROI for each subject using DPABI. The Pearson correlation analysis was performed to assess the relationship between the total GM volumes in each significant

<sup>1</sup><http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>

<sup>2</sup><http://dbm.neuro.uni-jena.de/cat/>

cluster and the magnitudes of vocal compensations for  $-50$  and  $-200$  cents perturbations with a significance level of  $p < 0.05$  for all statistical analyses.

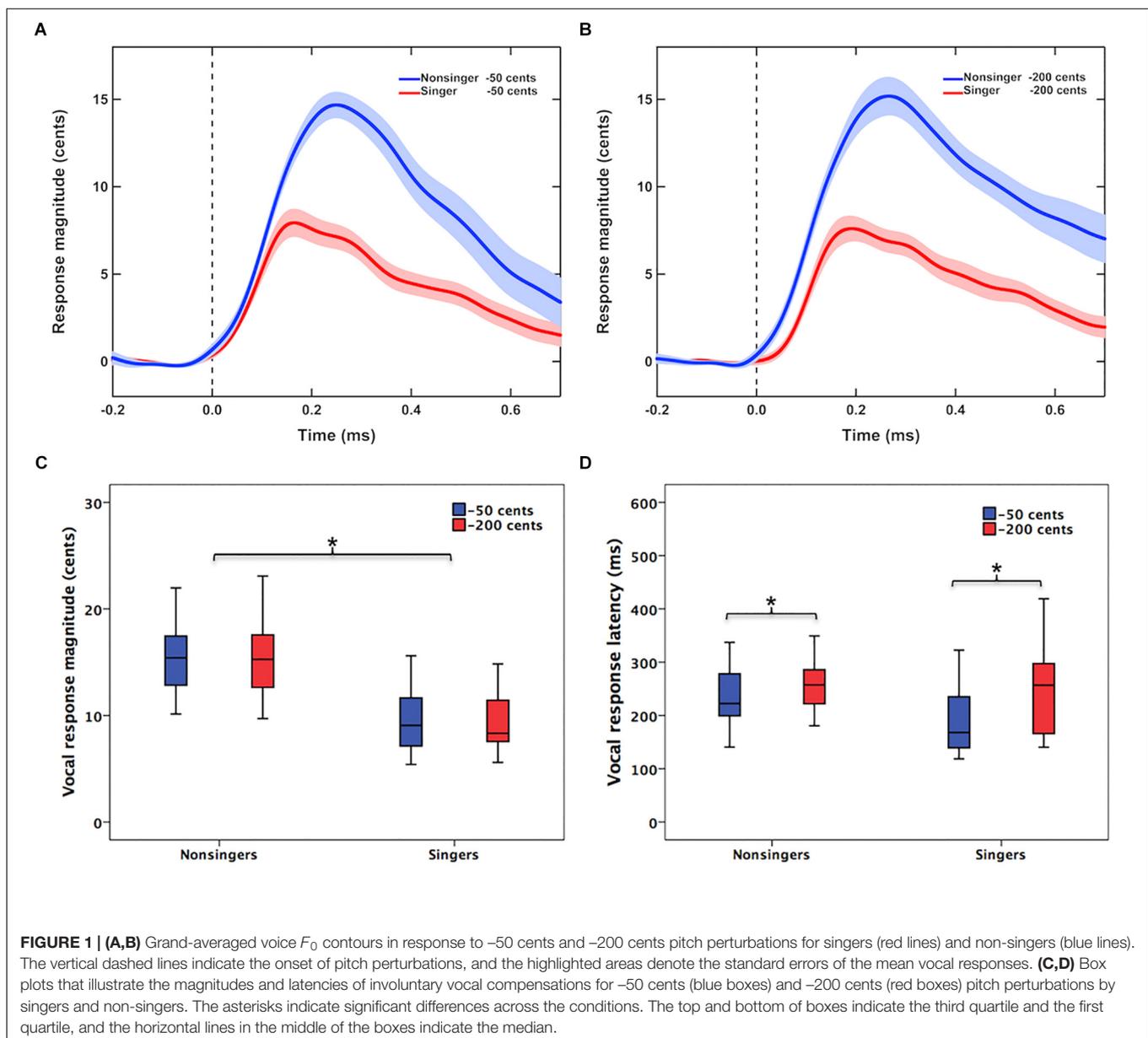
## RESULTS

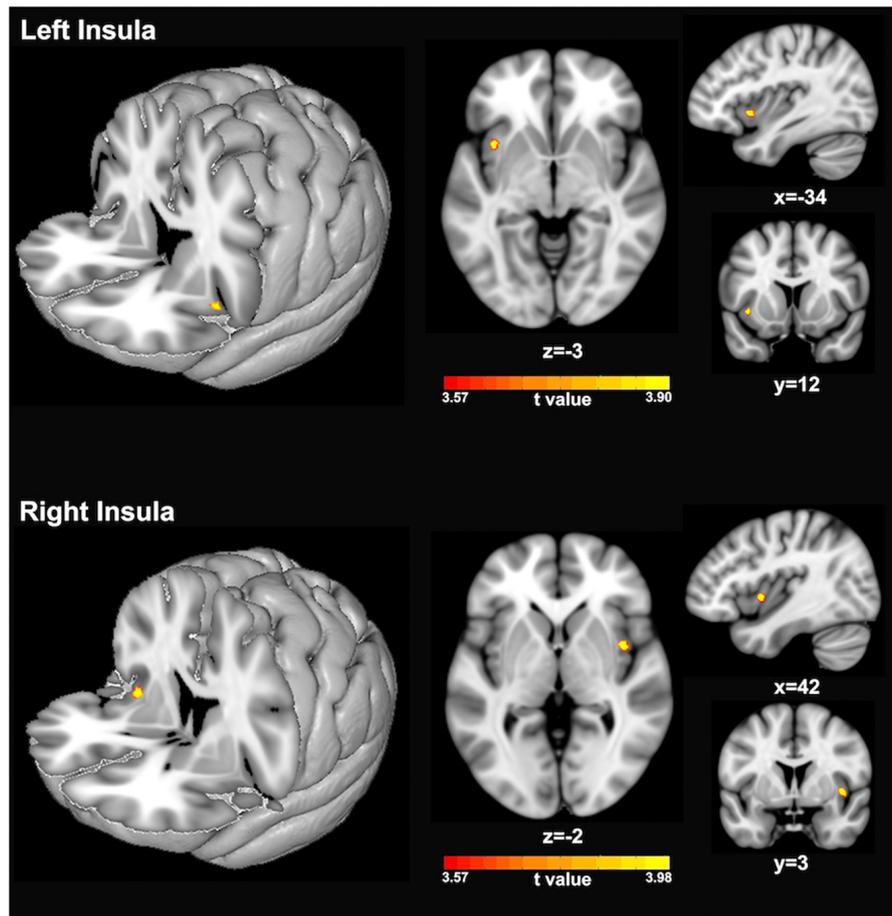
### Comparisons of Vocal Responses

**Figure 1** shows the grand-averaged voice  $F_0$  contours in response to  $-50$  and  $-200$  cents pitch perturbations for singers and non-singers. Regardless of the size of pitch perturbations, singers exhibited smaller involuntary vocal compensations than non-singers (**Figures 1A,B**). This difference was demonstrated by a significant main effect of group [ $F(1,40) = 36.834$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.479$ ] in a two-way RM-ANOVA conducted on the

magnitude of vocal compensation; singers produced significantly smaller vocal compensations than non-singers ( $9.4 \pm 2.7$  cents vs.  $15.8 \pm 4.4$  cents) (**Figure 1C**). The main effect of stimulus magnitude [ $F(1,40) = 0.003$ ,  $p = 0.953$ ,  $\eta_p^2 < 0.001$ ] ( $12.6 \pm 4.7$  cents vs.  $12.6 \pm 5.1$  cents), however, did not reach significance. Neither did the interaction between group and stimulus magnitude [ $F(1,40) = 0.656$ ,  $p = 0.423$ ,  $\eta_p^2 = 0.016$ ].

In addition, the latencies of vocal compensations were modulated as a function of stimulus magnitude [ $F(1,40) = 7.156$ ,  $p = 0.011$ ,  $\eta_p^2 = 0.152$ ], with faster vocal responses (i.e., smaller peak times) observed for  $-50$  cents perturbations than for  $-200$  cents perturbations ( $216 \pm 70$  vs.  $252 \pm 78$  ms) (**Figure 1D**). Although singers appeared to produce faster vocal responses than non-singers ( $223 \pm 90$  vs.  $244 \pm 57$  ms), this difference was not statistically significant [ $F(1,40) = 1.336$ ,  $p = 0.255$ ,  $\eta_p^2 = 0.032$ ].





**FIGURE 2 |** Statistical maps showing significantly lower GM volumes of left insula (**top panel**) and right insula (**bottom panel**) in singers compared to non-singers. The color bars indicate the *t*-values of two-sample *t*-test analyses.

The interaction between group and stimulus magnitude was not significant either [ $F(1,40) = 0.980$ ,  $p = 0.328$ ,  $\eta_p^2 = 0.024$ ].

### Comparisons of GM Volumes in Insula

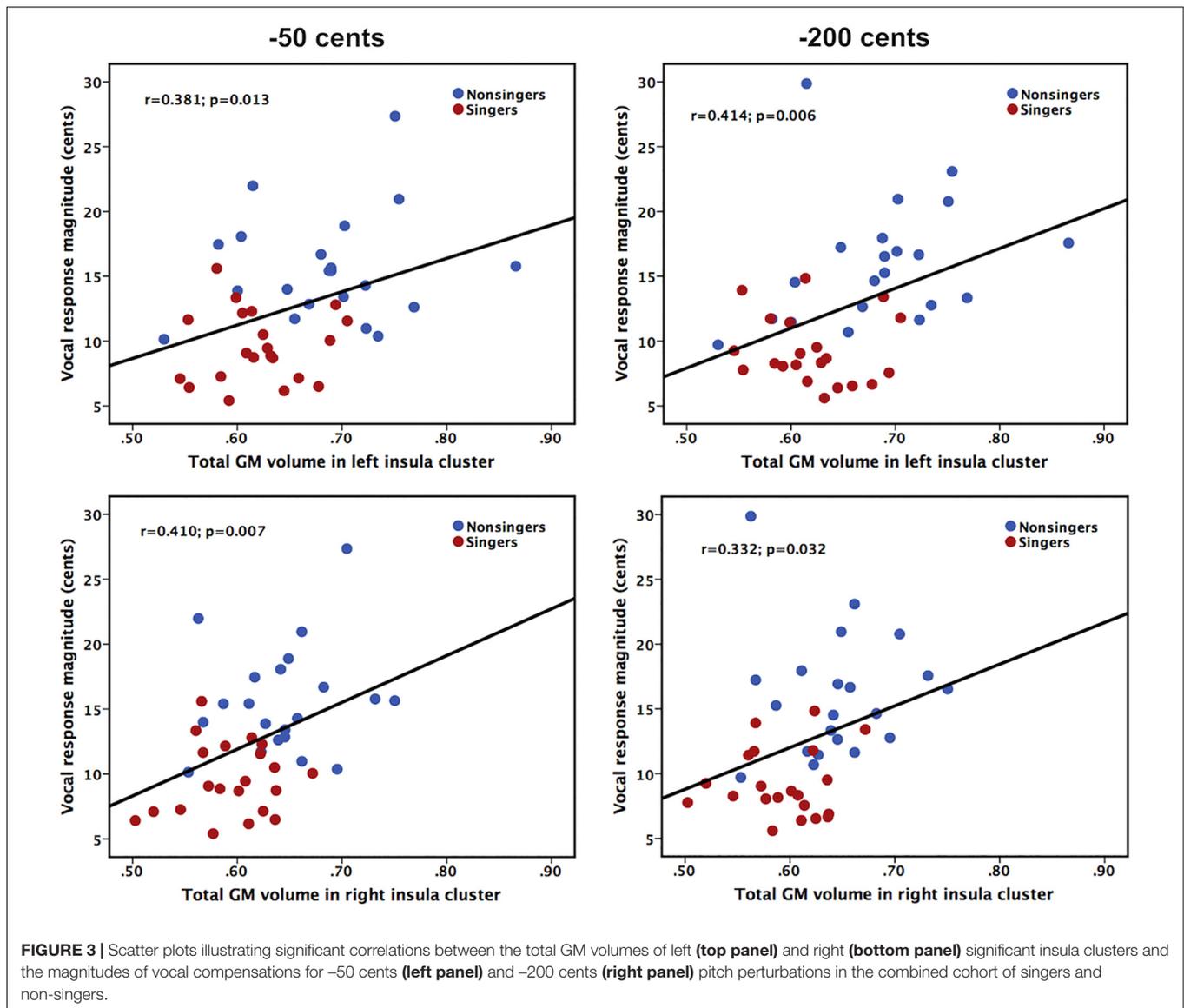
**Figure 2** and **Table 1** show the voxel-wise group comparison of GM volumes in left and right insula between singers and non-singers. As compared to non-singers, singers exhibited significantly lower GM volumes in the left insula (cluster level GRF corrected,  $p < 0.05$ ; MNI peak coordinate,  $-34\ 12\ -3$ ) and right insula (cluster level GRF corrected,  $p < 0.05$ , MIN peak coordinate,  $42\ 3\ -2$ ).

**TABLE 1 |** Brain regions that showed significantly smaller gray matter volumes in singers compared to non-singers.

| Cluster location       | MNI coordinates |    |    | Cluster size | t-value |
|------------------------|-----------------|----|----|--------------|---------|
|                        | x               | y  | z  |              |         |
| Left anterior insula   | -34             | 12 | -3 | 34           | 3.8965  |
| Right posterior insula | 42              | 3  | -2 | 45           | 3.9792  |

### The Brain–Behavior Relationship

In order to examine whether differences in insula morphology between singers and non-singers contributed to differences in their ability to compensate for voice pitch feedback perturbations, we performed Pearson correlation analyses by correlating the total GM volumes of significant insula clusters with the magnitudes of vocal compensations across the groups. As shown in **Figure 3**, on the combined cohort of both singers and non-singers, there were significant correlations between insula morphology and auditory-vocal integration. The total GM volume in the significant clusters of the left insula was positively correlated with the magnitude of vocal compensation for  $-50$  cents ( $r = 0.381$ ,  $p = 0.013$ ) and  $-200$  cents ( $r = 0.414$ ,  $p = 0.006$ ) pitch perturbations, respectively. Likewise, the total GM volume in the significant clusters of the right insula was positively correlated with the magnitude of vocal compensation for  $-50$  cents ( $r = 0.410$ ,  $p = 0.007$ ) and  $-200$  cents ( $r = 0.332$ ,  $p = 0.032$ ) pitch perturbations. Therefore, lower GM volumes in the left and right regional insula were predictive of smaller involuntary vocal compensations for pitch perturbations.



## DISCUSSION

The present study investigated structural changes in the insula induced by singing experience and their relationship with auditory feedback control of vocal pitch production. When perceiving pitch perturbations in voice auditory feedback, singers produced involuntary vocal compensations to a significantly lesser degree than non-singers. Interestingly, singers did not completely ignore larger pitch perturbations (i.e., -200 cents) but rather produced similar vocal responses to small pitch perturbations (i.e., -50 cents). VBM analysis revealed morphological changes in insula as a function of singing experience, as reflected by significantly lower GM volumes in the significant clusters of the left and right insula in singers relative to non-singers. Moreover, significant positive correlations were found between the total GM volumes of significant clusters in the insula and the magnitudes of compensatory vocal responses.

These findings provide morphometric evidence that reduced GM volume in insula contributes to singers' ability to suppress involuntary vocal responses to pitch perturbations during speaking as a consequence of singing experience.

### Speech Motor Control in Singers

Previous studies have shown that singers compensate for pitch perturbations to a lesser degree than non-singers in a voluntary manner when they sing the vowel/syllable at a specific note (Zarate and Zatorre, 2005, 2008; Jones and Keough, 2008). Likewise, our behavioral results revealed smaller involuntary vocal responses to pitch perturbations produced by singers compared to non-singers when they were instructed to vocalize the vowel sounds during speaking. In contrast, Behroozmand et al. (2014) reported no significant differences in the magnitudes of involuntary vocal compensations for pitch perturbations between musicians with AP and RP and non-musicians. We

speculate that differences between the population samples may be responsible for the inconsistency between our results and those of Behroozmand et al. (2014). For the present study we recruited a homogenous sample of female-only professional singers who did not possess AP, while Behroozmand et al. (2014) study involved both female and male singers and instrumentalists with AP and RP. Future studies need to be conducted to examine the differential effects of vocal and instrument training, as well as sex, on speech motor control.

Note that in the present study singers compensated for pitch perturbations to a lesser degree than non-singers regardless of the size of pitch perturbations, whereas in other studies (Zarate and Zatorre, 2008; Zarate et al., 2010) singers successfully ignored large pitch perturbations (e.g., 200 cents) but failed to ignore small perturbations (e.g., 25 cents). This disparity may be accounted for by methodological differences between the present study and Zarate and colleagues' work. The present study delivered a number of 200-ms-long pitch perturbations at mid-utterance and measured the vocal responses that began 100–300 ms after the perturbation onset. These early vocal responses are thought to be involuntary and cannot be consciously modulated by singers or non-singers (Munhall et al., 2009; Keough et al., 2013). In the studies by Zarate and Zatorre (2005, 2008) and Zarate et al. (2010), however, the pitch perturbations occurred 1–1.5 s after the vocal onset, lasted until the end of singing, and the vocal responses were measured only for the last second of each 4-s long singing. Their data, therefore, only included late vocal responses, which are subject to voluntary control in a top-down manner (Burnett et al., 1998; Hain et al., 2000; Patel et al., 2014). Thus, it is likely that singers are more successful than non-singers at suppressing vocal compensations for large pitch perturbations that are long enough to allow a voluntary response, but suppressed vocal compensations for both small and large perturbations in singers observed in the present study cannot be consciously modulated because these responses are involuntary in nature.

## Morphometric Changes in the Insula in Singers

Numerous studies have demonstrated that instrumental practice results in increased GM volume and cortical thickness in auditory, motor, and fronto-parietal regions as well as altered WM tracts (Schneider et al., 2002; Gaser and Schlaug, 2003; Bengtsson et al., 2005; Bermudez et al., 2009; Steele et al., 2013; Bailey et al., 2014; Groussard et al., 2014; Schlaug, 2015). For example, compared to non-musicians, instrumental musicians exhibited increased GM volume and cortical thickness in superior temporal regions (Bermudez et al., 2009), and increased brain activity in A1 and GM volume in the anteromedial portion of Heschl's gyrus (Schneider et al., 2002). Other research has shown a relationship between structural change and musical practice, with longer instrumental practice resulting in greater GM volumes in the left temporal and right frontal cortices, right somatosensory motor areas, and insula (Groussard et al., 2014). With regard to structural plasticity as a function of singing experience, Halwani et al. (2011) found a larger WM tract

volume in the right AF in both trained singers and instrumental musicians compared to non-musicians, while Kleber et al. (2016) found increased GM volumes in the right auditory and somatosensory cortices in singers. In the present study, we found significantly lower regional GM volumes in the left and right insula in singers relative to non-singers, providing further evidence in support of structural changes in GM and WM in brain regions that are fundamental to both singing and speaking as a function of singing experience. It is noteworthy that we also chose several other regions that are functionally related to speaking and singing (e.g., ACC, PMC, STG, etc.) as ROIs for VBM analyses but failed to find significant group differences in the GM volumes of those regions after multiple comparison correction, strengthening the role of the insula as an important hub within the speaking/singing network.

Interestingly, despite the previously reported increased GM volume, cortical thickness, or WM tract volume observed in singers, our results showed reduced regional GM volume in the insula in singers. Likewise, several other studies reported that smaller GM volumes were associated with better motor or cognitive performance (Draganski et al., 2006; Hanggi et al., 2010; Duan et al., 2012). Despite the methodological differences, our finding is accordance with a few studies showing negative relations between GM density/cortical thickness and speech learning proficiency (Beal et al., 2007; Dickerson et al., 2008; Rodriguez et al., 2018). For example, higher GM density in the IFG and STG and WM density in the insula were associated with adults who stutter relative to non-stutters (Beal et al., 2007), suggesting a relationship between atypical structural development and deficits in fluent speech production. Better performance on the California Verbal Learning Test was associated with thinner cortex in the paracentral/cingulate sulcus region (Dickerson et al., 2008). More recently, Rodriguez et al. (2018) found that thinner cortical thickness in the left anterior insula was predictive of better discrimination of novel speech sound contrasts in bilinguals, reflecting a more efficiently organized neural network that allows for increased speech learning proficiency. Similar effects of musical experience on the axonal membrane have been found, as reflected by lower fractional anisotropy values in the bilateral corticospinal tract in musicians relative to non-musicians (Schmithorst and Wilke, 2002; Imfeld et al., 2009). Although the neurobiology of structural plasticity in the brain remains unclear, the negative relationship between the volume of certain brain regions and speech and music performance may reflect more efficient neural organization caused by extensive experience. In light of this point, the reduced GM volume in the insula of the singers observed in the present study may reflect a refined neural network shaped by vocal training that allows for increased efficiency in the online detection and correction of errors in vocal output.

More importantly, we found that lower regional GM volumes in the bilateral insula were predictive of smaller vocal compensations for pitch perturbations. This finding is in line with the previously observed activation of insula in non-singers during the production of the early involuntary vocal compensations for pitch or F1 feedback perturbations during speaking (Toyomura et al., 2007; Tourville et al., 2008;

Parkinson et al., 2012; Behroozmand et al., 2015). As well, singers and non-singers exhibited increased activity in the insula when they were instructed to voluntarily ignore or compensate for pitch feedback perturbations during singing (Zarate and Zatorre, 2008; Zarate et al., 2010). Thus, our results not only confirm and extend previous findings that the insula plays a special role in the online control of speaking and singing but also provide evidence for linking morphometric differences in insula between singers and non-singers to their distinct behavioral performance in auditory feedback control of vocal production. In light of the lesion findings that showed abnormally increased vocal compensations in patients with Parkinson's disease, Alzheimer's disease, and cerebellar degeneration as a result of an overreliance on auditory feedback (Chen et al., 2013; Huang et al., 2016; Parrell et al., 2017; Ranasinghe et al., 2017), the observed association between lower GM volumes in the insula and smaller vocal compensations suggests that extensive vocal training may refine the insula-based networks to weigh less heavily on auditory feedback (see more details below), facilitating the precise control of vocal pitch production.

## Neural Mechanisms of Speech Motor Control in Singers

While most studies have focused on the neurobehavioral correlates of vocal pitch regulation during singing in singers (Burnett and Larson, 2002; Zarate and Zatorre, 2005, 2008; Jones and Keough, 2008; Keough and Jones, 2009; Keough et al., 2013), the present study examined the experience-dependent modulation of auditory-vocal integration using a speaking task. Despite the methodological differences across these studies, a consistent finding is the significantly lower degree to which singers compensate for pitch perturbations in voice auditory feedback relative to non-singers. Moreover, neuroimaging studies have shown a largely overlapping neural network that supports auditory-vocal integration during singing (Zarate and Zatorre, 2005, 2008; Zarate et al., 2010) and speaking (Tourville et al., 2008; Parkinson et al., 2012; Chang et al., 2013; Behroozmand et al., 2015, 2017; Guo et al., 2016). Therefore, both speaking and singing studies point to experience-dependent mechanisms that support sensorimotor control of vocal production.

A plausible explanation is that singers may rely less on auditory feedback and more on somatosensory feedback during vocal pitch regulation. According to the DIVA model (Golfinopoulos et al., 2010), auditory feedback and somatosensory feedback are closely correlated and tightly integrated for fine-tuning of vocal motor production. Although how these two types of feedback are integrated is currently unclear, a growing body of literature has focused on the role of somatosensory feedback in speech motor control. For example, an increased reliance on auditory feedback is observed when somatosensory feedback is absent, as evidenced by significantly increased vocal compensations for pitch perturbations after anesthetizing the vocal folds (Larson et al., 2008). Some individuals are even able to adapt to changes in voice auditory feedback with a heavy reliance on somatosensory feedback, as

reflected by the finding that participants who failed to adapt to auditory perturbations adapted to somatosensory perturbations when both auditory and somatosensory perturbations were presented simultaneously (Lametti et al., 2012). These findings suggest that a dynamic balance may exist between auditory and somatosensory feedback and when one form of feedback is compromised the speech motor system can compensate by using information from the other. In the context of singing, accumulating evidence has suggested that singers may weight somatosensory feedback more heavily for precise control of song production. For example, opera singers possessed increased GM volume in right S1 and S2 (Kleber et al., 2016) and exhibited increased activation in bilateral S1 during singing (Kleber et al., 2010). When auditory feedback was masked by noise, singers were still able to maintain pitch matching accuracy and showed increased activation in right anterior insula and its connectivity with SMG, whereas non-singers showed reduced pitch matching accuracy and decreased activity in right anterior insula and its connectivity with sensorimotor regions (Kleber et al., 2017). Therefore, it is possible that singers may develop a stronger reliance on somatosensory feedback to facilitate kinesthetic motor control of the vocal tract for the production of speech and song, thereby they are more capable of ignoring perturbations in auditory feedback and compensating for pitch perturbations to a lesser degree than non-singers.

In addition to feedback control, feedforward control also plays a critical role in speech motor control (Golfinopoulos et al., 2010) by enabling speakers to produce speech targets using previously learned motor commands without reliance on sensory feedback. There is evidence suggesting that singing expertise may lead to enhanced feedforward models that allow for precise control of vocal motor behaviors (Jones and Keough, 2008; Kleber et al., 2013). For example, Kleber et al. (2013) found that anesthesia of the vocal tract reduced pitch matching accuracy in singers to a lesser degree compared to non-singers, and this difference was accompanied by distinct pattern of brain activity under anesthesia across the two groups. Functional connectivity between right anterior insula, S1, A1, and M1 increased in non-singers, suggesting their recruitment of both the auditory and somatosensory network to produce the pitch targets. In contrast, singers exhibited decreased functional connectivity between the same areas and decreased activity in right anterior insula that was predictive of greater success in maintaining pitch matching accuracy under anesthesia (Kleber et al., 2013). This pattern of results suggests that the lack of somatosensory feedback did not lead to an increased reliance on auditory feedback; rather, singers may have largely relied on feedforward control developed during their vocal training that allowed them to ignore sensory feedback while maintaining pitch matching accuracy. This hypothesis is supported by one sensorimotor adaptation study by Jones and Keough (2008) that found higher voice  $F_0$  values during the aftereffects phase than those during the baseline phase for singers but not for non-singers. As an alternative explanation, therefore, singers may rely more heavily on feedforward control mechanisms but less on sensory feedback as a consequence of singing

experience to produce the desired vocal targets, which results in decreased vocal compensations for pitch perturbations in auditory feedback accordingly.

## Limitations

Several limitations of the present study should be acknowledged. One primary limitation is that the use of the AAL template restricted our ability to link vocal pitch regulation to insula morphology at the subregional level. The insula can be functionally divided into three subdivisions: the posterior, ventral anterior, and dorsal anterior insula (Deen et al., 2011; Uddin et al., 2014). Using probabilistic diffusion tractography, Battistella et al. (2018) revealed distinct subdivisions of the insula and parallel, largely non-overlapping WM pathways with cortical regions involved in different motor aspects of speech production ranging from articulatory modulations to communicative motivation. It is possible that different insula subregions may be differentially influenced by singing experience that may lead to different structural changes (e.g., increased or decreased GM volumes). Future studies should be conducted to examine the contributions of different insula subdivisions to the different patterns of vocal pitch regulation observed between singers and non-singers. Another limitation of the present study is that only female singers without AP were recruited. Previous studies have demonstrated sex and AP effects on brain morphology in musicians (Hutchinson et al., 2003; Lee et al., 2003; Bermudez et al., 2009). There is also evidence that shows that the behavioral and neural processing of vocal pitch regulation varies as a function of sex (Chen et al., 2010; Swink and Stuart, 2012; Li et al., 2018) and AP (Behroozmand et al., 2014). Therefore, our results may not generalize across musician populations of both sexes or to individuals with AP. Finally, it has been documented that auditory feedback control of speech production varies as a function of language experience (e.g., Mandarin, Cantonese, English) (Liu et al., 2010; Chen et al., 2012; Ning et al., 2014, 2015). Therefore, we cannot also rule out the possibility that the relationship between insula morphology and vocal pitch regulation differs across different languages.

## REFERENCES

- Ackermann, H., and Riecker, A. (2010). The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Struct. Funct.* 214, 419–433. doi: 10.1007/s00429-010-0257-x
- Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry—the methods. *NeuroImage* 11(6 Pt 1), 805–821. doi: 10.1006/nimg.2000.0582
- Bailey, J. A., Zatorre, R. J., and Penhune, V. B. (2014). Early musical training is linked to gray matter structure in the ventral premotor cortex and auditory-motor rhythm synchronization performance. *J. Cogn. Neurosci.* 26, 755–767. doi: 10.1162/jocn\_a\_00527
- Bamiou, D. E., Musiek, F. E., Stow, I., Stevens, J., Cipolotti, L., Brown, M. M., et al. (2006). Auditory temporal processing deficits in patients with insular stroke. *Neurology* 67, 614–619. doi: 10.1212/01.wnl.0000230197.40410.db
- Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., et al. (2003). Voxel-based lesion-symptom mapping. *Nat. Neurosci.* 6, 448–450.
- Battistella, G., Kumar, V., and Simonyan, K. (2018). Connectivity profiles of the insular network for speech control in healthy individuals and patients

## CONCLUSION

The present study investigated the association between structural plasticity in insula as a function of singing experience and auditory-motor integration during vocal pitch regulation. Singers showed reduced regional GM volumes in the bilateral insula and decreased involuntary vocal compensations for pitch perturbations than non-singers, and smaller regional GM volumes in insula were significantly correlated with the magnitudes of vocal compensations. It is suggested that differential feedback and feedforward mechanisms may underlie the distinct pattern of speech motor control between singers and non-singers, which may be related to morphometric changes in insula as a result of singing training.

## ETHICS STATEMENT

All participants gave written informed consent in compliance with a research protocol approved by the Institutional Review Board of The First Affiliated Hospital at Sun Yat-sen University of China.

## AUTHOR CONTRIBUTIONS

WW, GG, and HL designed the experiments. WW, LW, and NC performed the experiments and analyzed the data. WW, JJ, GG, and HL interpreted the results and wrote the manuscript. All authors read and approved the final manuscript.

## FUNDING

This study was funded by grants from the National Natural Science Foundation of China (Nos. 31371135, 81472154, and 81772439), the Guangdong Province Science and Technology Planning Project (No. 2017A050501014), and the Guangzhou Science and Technology Program (No. 201604020115).

with spasmodic dysphonia. *Brain Struct. Funct.* 223, 2489–2498. doi: 10.1007/s00429-018-1644-y

- Bauer, J. J., Mittal, J., Larson, C. R., and Hain, T. C. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: an automatic mechanism for stabilizing voice amplitude. *J. Acoust. Soc. Am.* 119, 2363–2371. doi: 10.1121/1.2173513
- Beal, D. S., Gracco, V. L., Lafaille, S. J., and De Nil, L. F. (2007). Voxel-based morphometry of auditory and speech-related cortex in stutterers. *Neuroreport* 18, 1257–1260. doi: 10.1097/wnr.0b013e3282202c4d
- Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A., and Larson, C. R. (2014). Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain Cogn.* 84, 97–108. doi: 10.1016/j.bandc.2013.11.007
- Behroozmand, R., Phillip, L., Johari, K., Bonilha, L., Rorden, C., Hickok, G., et al. (2017). Sensorimotor impairment of speech auditory feedback processing in aphasia. *NeuroImage* 165, 102–111. doi: 10.1016/j.neuroimage.2017.10.014
- Behroozmand, R., Shebek, R., Hansen, D. R., Oya, H., Robin, D. A., Howard, M. A. III, et al. (2015). Sensory-motor networks involved in speech production

- and motor control: an fMRI study. *NeuroImage* 109, 418–428. doi: 10.1016/j.neuroimage.2015.01.040
- Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., and Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* 8, 1148–1150. doi: 10.1038/nn1516
- Bermudez, P., Lerch, J. P., Evans, A. C., and Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cereb. Cortex* 19, 1583–1596. doi: 10.1093/cercor/bhn196
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott Int.* 5, 341–345.
- Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *J. Acoust. Soc. Am.* 103, 3153–3161. doi: 10.1121/1.423073
- Burnett, T. A., and Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *J. Acoust. Soc. Am.* 112, 1058–1063. doi: 10.1121/1.1487844
- Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., and Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2653–2658. doi: 10.1073/pnas.1216827110
- Chen, S. H., Liu, H., Xu, Y., and Larson, C. R. (2007). Voice F0 responses to pitch-shifted voice feedback during english speech. *J. Acoust. Soc. Am.* 121, 1157–1163. doi: 10.1121/1.2404624
- Chen, X., Zhu, X., Wang, E. Q., Chen, L., Li, W., Chen, Z., et al. (2013). Sensorimotor control of vocal pitch production in Parkinson's disease. *Brain Res.* 1527, 99–107. doi: 10.1016/j.brainres.2013.06.030
- Chen, Z., Liu, P., Jones, J. A., Huang, D., and Liu, H. (2010). Sex-related differences in vocal responses to pitch feedback perturbations during sustained vocalization. *J. Acoust. Soc. Am.* 128, EL355–EL360. doi: 10.1121/1.3509124
- Chen, Z., Liu, P., Wang, E. Q., Larson, C. R., Huang, D., and Liu, H. (2012). ERP correlates of language-specific processing of auditory pitch feedback during self-vocalization. *Brain Lang.* 121, 25–34. doi: 10.1016/j.bandl.2012.02.004
- Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., and Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195. doi: 10.1038/nn1176
- Deen, B., Pitskel, N. B., and Pelphrey, K. A. (2011). Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* 21, 1498–1506. doi: 10.1093/cercor/bhq186
- Dickerson, B. C., Fenstermacher, E., Salat, D. H., Wolk, D. A., Maguire, R. P., Desikan, R., et al. (2008). Detection of cortical thickness correlates of cognitive performance: reliability across MRI scan sessions, scanners, and field strengths. *NeuroImage* 39, 10–18. doi: 10.1016/j.neuroimage.2007.08.042
- Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Buchel, C., et al. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning. *J. Neurosci.* 26, 6314–6317. doi: 10.1523/jneurosci.4628-05.2006
- Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature* 384, 159–161. doi: 10.1038/384159a0
- Duan, X., He, S., Liao, W., Liang, D., Qiu, L., Wei, L., et al. (2012). Reduced caudate volume and enhanced striatal-DMN integration in chess experts. *NeuroImage* 60, 1280–1286. doi: 10.1016/j.neuroimage.2012.01.047
- Gaser, C., and Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245. doi: 10.1523/jneurosci.23-27-09240.2003
- Ghaziri, J., Tucholka, A., Girard, G., Houde, J. C., Boucher, O., Gilbert, G., et al. (2017). The corticocortical structural connectivity of the human insula. *Cereb. Cortex* 27, 1216–1228. doi: 10.1093/cercor/bhv308
- Glasser, M. F., and Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cereb. Cortex* 18, 2471–2482. doi: 10.1093/cercor/bhn011
- Golfinopoulos, E., Tourville, J. A., and Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage* 52, 862–874. doi: 10.1016/j.neuroimage.2009.10.023
- Grell, A., Sundberg, J., Ternstrom, S., Ptok, M., and Altenmuller, E. (2009). Rapid pitch correction in choir singers. *J. Acoust. Soc. Am.* 126, 407–413. doi: 10.1121/1.3147508
- Groussard, M., Viader, F., Landeau, B., Desgranges, B., Eustache, F., and Platel, H. (2014). The effects of musical practice on structural plasticity: the dynamics of grey matter changes. *Brain Cogn.* 90, 174–180. doi: 10.1016/j.bandc.2014.06.013
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *J. Commun. Disord.* 39, 350–365. doi: 10.1016/j.jcomdis.2006.06.013
- Guo, Z., Huang, X., Wang, M., Jones, J. A., Dai, Z., Li, W., et al. (2016). Regional homogeneity of intrinsic brain activity correlates with auditory-motor processing of vocal pitch errors. *NeuroImage* 142, 565–575. doi: 10.1016/j.neuroimage.2016.08.005
- Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., and Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Exp. Brain Res.* 130, 133–141. doi: 10.1007/s002219900237
- Halwani, G. F., Loui, P., Ruber, T., and Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Front. Psychol.* 2:156. doi: 10.3389/fpsyg.2011.00156
- Hanggi, J., Koeneke, S., Bezzola, L., and Jancke, L. (2010). Structural neuroplasticity in the sensorimotor network of professional female ballet dancers. *Hum. Brain Mapp.* 31, 1196–1206. doi: 10.1002/hbm.20928
- Houde, J. F., and Chang, E. F. (2015). The cortical computations underlying feedback control in vocal production. *Curr. Opin. Neurobiol.* 33, 174–181. doi: 10.1016/j.conb.2015.04.006
- Houde, J. F., and Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science* 279, 1213–1216. doi: 10.1126/science.279.5354.1213
- Huang, X., Chen, X., Yan, N., Jones, J. A., Wang, E. Q., Chen, L., et al. (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. doi: 10.1002/hbm.23306
- Hutchinson, S., Lee, L. H., Gaab, N., and Schlaug, G. (2003). Cerebellar volume of musicians. *Cereb. Cortex* 13, 943–949. doi: 10.1093/cercor/13.9.943
- Imfeld, A., Oechslin, M. S., Meyer, M., Loenneker, T., and Jancke, L. (2009). White matter plasticity in the corticospinal tract of musicians: a diffusion tensor imaging study. *NeuroImage* 46, 600–607. doi: 10.1016/j.neuroimage.2009.02.025
- Jones, J. A., and Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Exp. Brain Res.* 190, 279–287. doi: 10.1007/s00221-008-1473-y
- Jones, J. A., and Munhall, K. G. (2005). Remapping auditory-motor representations in voice production. *Curr. Biol.* 15, 1768–1772. doi: 10.1016/j.cub.2005.08.063
- Keough, D., Hawco, C., and Jones, J. A. (2013). Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback. *BMC Neurosci.* 14:25. doi: 10.1186/1471-2202-14-25
- Keough, D., and Jones, J. A. (2009). The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *J. Acoust. Soc. Am.* 126, 837–846. doi: 10.1121/1.3158600
- Kleber, B., Friberg, A., Zeitouni, A., and Zatorre, R. (2017). Experience-dependent modulation of right anterior insula and sensorimotor regions as a function of noise-masked auditory feedback in singers and nonsingers. *NeuroImage* 147, 97–110. doi: 10.1016/j.neuroimage.2016.11.059
- Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., and Lotze, M. (2010). The brain of opera singers: experience-dependent changes in functional activation. *Cereb. Cortex* 20, 1144–1152. doi: 10.1093/cercor/bhp177
- Kleber, B., Veit, R., Moll, C. V., Gaser, C., Birbaumer, N., and Lotze, M. (2016). Voxel-based morphometry in opera singers: increased gray-matter volume in right somatosensory and auditory cortices. *NeuroImage* 133, 477–483. doi: 10.1016/j.neuroimage.2016.03.045
- Kleber, B., Zeitouni, A. G., Friberg, A., and Zatorre, R. J. (2013). Experience-dependent modulation of feedback integration during singing: role of the right anterior insula. *J. Neurosci.* 33, 6070–6080. doi: 10.1523/JNEUROSCI.4418-12.2013
- Kort, N. S., Cuesta, P., Houde, J. F., and Nagarajan, S. S. (2016). Bihemispheric network dynamics coordinating vocal feedback control. *Hum. Brain Mapp.* 37, 1474–1485. doi: 10.1002/hbm.23114
- Kort, N. S., Nagarajan, S. S., and Houde, J. F. (2014). A bilateral cortical network responds to pitch perturbations in speech feedback. *NeuroImage* 86, 525–535. doi: 10.1016/j.neuroimage.2013.09.042
- Lametti, D. R., Nasir, S. M., and Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *J. Neurosci.* 32, 9351–9358. doi: 10.1523/JNEUROSCI.0404-12.2012

- Larson, C. R., Altman, K. W., Liu, H., and Hain, T. C. (2008). Interactions between auditory and somatosensory feedback for voice F (0) control. *Exp. Brain Res.* 187, 613–621. doi: 10.1007/s00221-008-1330-z
- Lee, D. J., Chen, Y., and Schlaug, G. (2003). Corpus callosum: musician and gender effects. *Neuroreport* 14, 205–209. doi: 10.1097/00001756-200302100-00009
- Li, J., Hu, H., Chen, N., Jones, J. A., Wu, D., Liu, P., et al. (2018). Aging and sex influence cortical auditory-motor integration for speech control. *Front. Neurosci.* 12:749. doi: 10.3389/fnins.2018.00749
- Li, W., Chen, Z., Liu, P., Zhang, B., Huang, D., and Liu, H. (2013). Neurophysiological evidence of differential mechanisms involved in producing opposing and following responses to altered auditory feedback. *Clin. Neurophysiol.* 124, 2161–2171. doi: 10.1016/j.clinph.2013.04.340
- Liu, H., and Larson, C. R. (2007). Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *J. Acoust. Soc. Am.* 122, 3671–3677. doi: 10.1121/1.2800254
- Liu, H., Meshman, M., Behroozmand, R., and Larson, C. R. (2011). Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clin. Neurophysiol.* 122, 951–957. doi: 10.1016/j.clinph.2010.08.010
- Liu, H., Wang, E. Q., Chen, Z., Liu, P., Larson, C. R., and 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. doi: 10.1121/1.3500675
- Munhall, K. G., MacDonald, E. N., Byrne, S. K., and Johnsrude, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *J. Acoust. Soc. Am.* 125, 384–390. doi: 10.1121/1.3035829
- Murbe, D., Pabst, F., Hofmann, G., and Sundberg, J. (2004). Effects of a professional solo singer education on auditory and kinesthetic feedback—a longitudinal study of singers' pitch control. *J. Voice* 18, 236–241. doi: 10.1016/j.jvoice.2003.05.001
- Natke, U., Donath, T. M., and Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *J. Acoust. Soc. Am.* 113, 1587–1593. doi: 10.1121/1.1543928
- Ning, L., Loucks, T. M., and Shih, C. (2015). The effects of language learning and vocal training on sensorimotor control of lexical tone. *J. Phon.* 51, 50–69. doi: 10.1016/j.wocn.2014.12.003
- Ning, L., Shih, C., and Loucks, T. M. (2014). Mandarin tone learning in L2 adults: a test of perceptual and sensorimotor contributions. *Speech Commun.* 63–64, 55–69. doi: 10.1016/j.specom.2014.05.001
- Ozdemir, E., Norton, A., and Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *NeuroImage* 33, 628–635. doi: 10.1016/j.neuroimage.2006.07.013
- Parkinson, A. L., Flagmeier, S. G., Manes, J. L., Larson, C. R., Rogers, B., and Robin, D. A. (2012). Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage* 61, 314–322. doi: 10.1016/j.neuroimage.2012.02.068
- Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., and Ivry, R. B. (2017). Impaired feedforward control and enhanced feedback control of speech in patients with cerebellar degeneration. *J. Neurosci.* 37, 9249–9258. doi: 10.1523/JNEUROSCI.3363-16.2017
- Patel, S., Nishimura, C., Lodhavia, A., Korzyukov, O., Parkinson, A., Robin, D. A., et al. (2014). Understanding the mechanisms underlying voluntary responses to pitch-shifted auditory feedback. *J. Acoust. Soc. Am.* 135, 3036–3044. doi: 10.1121/1.4870490
- Purcell, D. W., and Munhall, K. G. (2006). Compensation following real-time manipulation of formants in isolated vowels. *J. Acoust. Soc. Am.* 119, 2288–2297. doi: 10.1121/1.2173514
- Ranasinghe, K. G., Gill, J. S., Kothare, H., Beagle, A. J., Mizuir, D., Honma, S. M., et al. (2017). Abnormal vocal behavior predicts executive and memory deficits in Alzheimer's disease. *Neurobiol. Aging* 52, 71–80. doi: 10.1016/j.neurobiolaging.2016.12.020
- Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., and Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport* 11, 1997–2000. doi: 10.1097/00001756-200006260-00038
- Riecker, A., Mathiak, K., Wildgruber, D., Erb, M., Hertrich, I., Grodd, W., et al. (2005). fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology* 64, 700–706. doi: 10.1212/01.wnl.0000152156.90779.89
- Rodriguez, S. M., Archila-Suerte, P., Vaughn, K. A., Chiarello, C., and Hernandez, A. E. (2018). Anterior insular thickness predicts speech sound learning ability in bilinguals. *NeuroImage* 165, 278–284. doi: 10.1016/j.neuroimage.2017.10.038
- Scheerer, N. E., Behich, J., Liu, H., and Jones, J. A. (2013). ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience* 240, 176–185. doi: 10.1016/j.neuroscience.2013.02.054
- Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. *Prog. Brain Res.* 217, 37–55. doi: 10.1016/bs.pbr.2014.11.020
- Schmithorst, V. J., and Wilke, M. (2002). Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neurosci. Lett.* 321, 57–60. doi: 10.1016/s0304-3940(02)00054-x
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., and Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat. Neurosci.* 5, 688–694. doi: 10.1038/nn871
- Shelley, B. P., and Trimble, M. R. (2004). The insular lobe of Reil—its anatomic-functional, behavioural and neuropsychiatric attributes in humans—a review. *World J Biol Psychiatry* 5, 176–200. doi: 10.1080/15622970410029933
- Steele, C. J., Bailey, J. A., Zatorre, R. J., and Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: evidence for a sensitive period. *J. Neurosci.* 33, 1282–1290. doi: 10.1523/JNEUROSCI.3578-12.2013
- Swink, S., and Stuart, A. (2012). The effect of gender on the N1-P2 auditory complex while listening and speaking with altered auditory feedback. *Brain Lang.* 122, 25–33. doi: 10.1016/j.bandl.2012.04.007
- Tang, X., Chen, N., Zhang, S., Jones, J. A., Zhang, B., Li, J., et al. (2018). Predicting auditory feedback control of speech production from subregional shape of subcortical structures. *Hum. Brain Mapp.* 39, 459–471. doi: 10.1002/hbm.23855
- Tourville, J. A., Reilly, K. J., and Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage* 39, 1429–1443. doi: 10.1016/j.neuroimage.2007.09.054
- Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., et al. (2007). Neural correlates of auditory feedback control in human. *Neuroscience* 146, 499–503. doi: 10.1016/j.neuroscience.2007.02.023
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Uddin, L. Q., Kinnison, J., Pessoa, L., and Anderson, M. L. (2014). Beyond the tripartite cognition-emotion-interoception model of the human insular cortex. *J. Cogn. Neurosci.* 26, 16–27. doi: 10.1162/jocn\_a\_00462
- Yan, C. G., Wang, X. D., Zuo, X. N., and Zang, Y. F. (2016). DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/s12021-016-9299-4
- Zarate, J. M., Wood, S., and Zatorre, R. J. (2010). Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia* 48, 607–618. doi: 10.1016/j.neuropsychologia.2009.10.025
- Zarate, J. M., and Zatorre, R. J. (2005). Neural substrates governing audiovisual integration for vocal pitch regulation in singing. *Ann. N. Y. Acad. Sci.* 1060, 404–408. doi: 10.1196/annals.1360.058
- Zarate, J. M., and Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage* 40, 1871–1887. doi: 10.1016/j.neuroimage.2008.01.026
- Zatorre, R. J., and Baum, S. R. (2012). Musical melody and speech intonation: singing a different tune. *PLoS Biol.* 10:e1001372. doi: 10.1371/journal.pbio.1001372

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Wang, Wei, Chen, Jones, Gong and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.