



# Auditory-Motor Control of Vocal Production during Divided Attention: Behavioral and ERP Correlates

Ying Liu<sup>1†</sup>, Hao Fan<sup>1†</sup>, Jingting Li<sup>1†</sup>, Jeffery A. Jones<sup>2</sup>, Peng Liu<sup>1</sup>, Baofeng Zhang<sup>1\*</sup> and Hanjun Liu<sup>1,3\*</sup>

<sup>1</sup> Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, <sup>2</sup> Psychology Department and Laurier Centre for Cognitive Neuroscience, Wilfrid Laurier University, Waterloo, ON, Canada, <sup>3</sup> Guangdong Provincial Key Laboratory of Brain Function and Disease, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China

## OPEN ACCESS

### Edited by:

Huan Luo,  
Peking University, China

### Reviewed by:

Dan Zhang,  
Tsinghua University, China  
Soo-Eun Chang,  
University of Michigan Health System,  
United States

### \*Correspondence:

Hanjun Liu  
lhanjun@mail.sysu.edu.cn  
Baofeng Zhang  
long.sums@163.com

<sup>†</sup>These authors have contributed equally to this work and shared first authorship.

### Specialty section:

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

**Received:** 18 November 2017

**Accepted:** 13 February 2018

**Published:** 27 February 2018

### Citation:

Liu Y, Fan H, Li J, Jones JA, Liu P,  
Zhang B and Liu H (2018)  
Auditory-Motor Control of Vocal  
Production during Divided Attention:  
Behavioral and ERP Correlates.  
*Front. Neurosci.* 12:113.  
doi: 10.3389/fnins.2018.00113

When people hear unexpected perturbations in auditory feedback, they produce rapid compensatory adjustments of their vocal behavior. Recent evidence has shown enhanced vocal compensations and cortical event-related potentials (ERPs) in response to attended pitch feedback perturbations, suggesting that this reflex-like behavior is influenced by selective attention. Less is known, however, about auditory-motor integration for voice control during divided attention. The present cross-modal study investigated the behavioral and ERP correlates of auditory feedback control of vocal pitch production during divided attention. During the production of sustained vowels, 32 young adults were instructed to simultaneously attend to both pitch feedback perturbations they heard and flashing red lights they saw. The presentation rate of the visual stimuli was varied to produce a low, intermediate, and high attentional load. The behavioral results showed that the low-load condition elicited significantly smaller vocal compensations for pitch perturbations than the intermediate-load and high-load conditions. As well, the cortical processing of vocal pitch feedback was also modulated as a function of divided attention. When compared to the low-load and intermediate-load conditions, the high-load condition elicited significantly larger N1 responses and smaller P2 responses to pitch perturbations. These findings provide the first neurobehavioral evidence that divided attention can modulate auditory feedback control of vocal pitch production.

**Keywords:** auditory feedback, speech motor control, divided attention, attentional load, working memory

## INTRODUCTION

Auditory feedback is critical for the production of proper speech sounds (Hickok et al., 2011). Numerous behavioral studies have demonstrated that speakers compensate for alterations in voice pitch, loudness, and formant frequencies by producing vocal adjustments against the direction of the alterations (Burnett et al., 1998; Jones and Munhall, 2002; Bauer et al., 2006; Purcell and Munhall, 2006; Liu and Larson, 2007; Liu et al., 2007; Macdonald et al., 2010; Mitsuya et al., 2015). This compensatory control process can be modulated by task demands (Natke et al., 2003; Chen et al., 2007) and shaped by language and music experience (Zarate and Zatorre, 2008; Liu et al., 2010; Mitsuya et al., 2013; Behroozmand et al., 2014). Furthermore, patients with Parkinson's disease (PD), Alzheimer's disease (AD), and temporal lobe epilepsy (TLE) produce abnormally

enhanced vocal compensations for pitch perturbations (Liu et al., 2012; Chen et al., 2013; Mollaei et al., 2016; Ranasinghe et al., 2017). Thus, understanding the mechanisms that underlie auditory feedback control of speech production is important for the treatment of voice/speech disorders caused by neurological diseases.

Brief auditory feedback perturbations typically evoke rapid compensatory vocal responses with short latencies of ~80–150 ms (Larson, 1998; Burnett and Larson, 2002; Chen et al., 2007; Liu and Larson, 2007). Moreover, speakers are unable to consciously modify their vocal compensations even when told to, suggesting that the feedback-based control of speech production is a reflex-like process (Munhall et al., 2009; Keough et al., 2013). There is evidence, however, suggesting that auditory feedback control of speech production may be subject to attentional control. Previous research has repeatedly shown that attended auditory stimuli elicit larger event-related potentials (ERPs) (Hink and Hillyard, 1976; Stevens et al., 2006) and enhanced brain activity in the auditory cortex (Ahveninen et al., 2006; Johnson and Zatorre, 2006; Sabri et al., 2008) relative to unattended auditory stimuli. These findings suggest that the perception of speech sounds is highly dependent on attention. Similarly, auditory-motor interactions during speech processing can also be facilitated by attention, as reflected by increased left-hemisphere P50m responses when participants attended to lip-articulated “ba” sounds while their cortical motor lip area was disrupted by transcranial magnetic stimulation (TMS) (Möttönen et al., 2014). In the context of speech motor control, Tumber et al. (2014) reported that participants who were exposed to pitch perturbations during vocalization produced smaller vocal compensations when they actively attended to a rapid serial visual presentation (RSVP) of letters relative to when they passively viewed the RSVP, suggesting that the attentional load of the RSVP task reduced the available attentional resources for the detection and/or correction for production errors. In two other studies conducted in our laboratory (Hu et al., 2015; Liu et al., 2015), participants were asked to attend to pitch perturbations they heard in voice auditory feedback, or attend to flashing lights they viewed during the production of sustained vowels. The results showed that attending to pitch perturbations elicited significantly larger vocal compensations and P2 responses relative to ignoring pitch perturbations (i.e., attending to flashing lights) and passively observing the bimodal stimuli. These neurobehavioral findings can be accounted for by the gain-based theory of selective attention (Hillyard et al., 1998), according to which selective attention increases the gain for neurons involved in auditory-vocal integration which in turn facilitates the detection/correction of voice feedback errors.

It is noteworthy that during daily communication attention is often divided such that auditory feedback can be processed in conjunction with other sensory information simultaneously. The effect of divided attention on auditory-motor integration for voice control, however, is far from clear. Previous studies have shown decreased brain activity in both the auditory and visual cortices but increased brain activity in the lateral frontal regions when attention is divided between auditory and visual stimuli compared to when attention is focused on either auditory

or visual stimuli alone (Klingberg, 1998; Loose et al., 2003; Johnson and Zatorre, 2006; Moisaala et al., 2015). In one recent ERP study by Getzmann et al. (2016), dividing attention to speech from two speakers led to smaller N1 and P2 responses than focusing attention on speech from one speaker. Likewise, this dual-task interference during divided attention influences auditory-motor control of vocal production. Liu et al. (2015) reported significantly smaller P2 responses when attention was divided to both pitch perturbation and flashing lights as compared to when pitch perturbations were selectively attended and ignored. In addition, dividing attention to the bimodal stimuli elicited significantly larger N1 responses and smaller P2 responses relative to passively observing the bimodal stimuli. These findings suggest that divided attention can modulate the cortical processing of mismatches between intended and actual vocal output.

Certain shortcomings in the study by Liu et al. (2015), however, limit our understanding of how divided attention influences auditory-motor control of vocal production. First of all, although the N1 and P2 responses have been hypothesized to, respectively, reflect the early detection of mismatches between predicted and actual voice auditory feedback and the later cortical activity involved in auditory-motor interaction (Behroozmand et al., 2011; Guo et al., 2016), the observed modulation of N1 and P2 responses in Liu et al. (2015) may instead be the result of attention-driven central auditory processing of pitch feedback errors because vocal compensations did not vary as a function of divided attention. Next, as compared to when participants divided attention to pitch perturbations and flashing lights, N1 responses were significantly larger when participants passively observed the bimodal stimuli whereas remained intact when participants attended to flashing lights while ignoring pitch perturbations. Whether this N1 enhancement was a result of divided attention remains unclear. Finally, given that the lateral frontal regions subserving working memory were recruited during divided attention but not selective attention (Johnson and Zatorre, 2006), Liu et al. (2015) attributed the modulation of cortical N1 and P2 responses to pitch perturbations during divided attention to the interaction between working memory and divided attention. There is at present insufficient evidence, however, to support this hypothesis. Thus, the present study aims to extend results from our previous investigation (Liu et al., 2015) and thereby expand current knowledge about the interaction between divided attention and auditory-vocal integration.

In summary, little is currently known about the effect of divided attention on the auditory-motor control of speech production. In order to address this important question, the present study examined the behavioral and ERP correlates of auditory feedback-based vocal pitch regulation during divided attention. We adapted the previously used paradigm (Liu et al., 2015), during which participants were instructed to attend to pitch perturbations in auditory feedback and red indicator lights on the screen simultaneously while producing sustained vowels. These two sensory stimuli were behaviorally irrelevant and their presentation did not overlap to avoid any interaction between multisensory integration and attention control. In order to vary the attentional resources available for auditory feedback

processing during divided attention, we varied the inter-stimulus intervals (ISIs) in the presentation rate of the red indicator lights to impose a low, intermediate, and high attentional load. This paradigm has been successfully used in previous studies of divided attention (Craig et al., 1996; Naveh-Benjamin et al., 2000; Uncapher and Rugg, 2005) and allowed us to compare the neurobehavioral responses to pitch perturbations across the three attentional load levels. In light of our previous findings (Hu et al., 2015; Liu et al., 2015), we hypothesized that divided attention would exert modulatory effects on the neurobehavioral responses to pitch feedback errors during vocal production and that such effects would change as a function of attentional load.

## MATERIALS AND METHODS

### Subjects

Forty native Mandarin-speaking young adults participated in the present study. Eight participants were excluded from the final data pool because of poor data quality. Thus, data from thirty-two participants [21 female and 11 male; mean age and standard deviation (SD):  $21.53 \pm 2.41$  years] entered the final statistical analyses. They were all right-handed, had normal or corrected-normal vision, and reported no history of hearing, speech, language, or neurological disorders. Hearing thresholds were screened at 25 dB HL for octave intervals of 500–4000 Hz. Written informed consent was obtained from all participants. The research protocol that was in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) was approved by the Institutional Review Board of The First Affiliated Hospital at Sun Yat-sen University of China.

### Apparatus

The experiment was conducted in a sound-attenuated booth, where participants' voice and electroencephalographic (EEG) signals were recorded. In order to partially mask the air-born and bone-conducted feedback, we calibrated the acoustic recording system so that the intensity of voice feedback the participant heard was 10 dB SPL higher than that of his/her voice output. During the experiment, participants' voice signals were transduced by a dynamic microphone (DM2200, Takstar Inc.) and sent to an Eventide Eclipse Harmonizer via a MOTU Ultralite Mk3 Firewire audio interface. A custom-developed Max/MSP software program (v.5.0 by Cycling 74) controlled the Harmonizer to pitch-shift the voice signals and sent them to an ICON NeoAmp headphone amplifier. The amplified pitch-shifted voices were presented to participants as attended auditory stimuli through insert earphones (ER1-14A, Etymotic Research Inc.). Two circles representing the blue and red indicator lights were generated by the Max/MSP software program and displayed on the computer screen. The blue indicator light was used to cue the start and end of vocalization, while the flashing red indicator lights were used as attended visual stimuli. Transistor-transistor logic (TTL) control pulses were also generated by this program to mark the onset of the pitch perturbations. The original and pitch-shifted voice signals as well as the TTL control pulses were sampled at 10 kHz by a PowerLab A/D converter (ML880, AD

Instruments) and recorded using LabChart software (v.7.0 by AD Instruments).

While recording the voice signals, we collected the EEG signals from 64 sites on the participant's scalp using a Geodesic Sensor Net (Electrical Geodesics Inc.). Scalp-recorded brain potentials were amplified by a Net Amps 300 amplifier that accepts scalp-electrode impedances up to 40–60 k $\Omega$  ( $Z_{in} \approx 200$  M $\Omega$ ; Electrical Geodesics Inc.), digitized at 1 kHz, and recorded using NetStation software (v. 4.5, Electrical Geodesics Inc.). The TTL control pulses that signaled the onset of the pitch perturbations were sent to the EEG recording system via a DIN synch cable. The EEG signals across all channels were referenced to the vertex (Cz) during the online recording. Electrode impedances were maintained at  $\leq 50$  k $\Omega$  for individual sensors (Ferree et al., 2001).

### Procedure

In the present study, participants were instructed to produce and maintain a steady vocalization of the vowel /u/ at their comfortable pitch and loudness level when the blue indicator light was turned on and terminate their vocalizations when the blue indicator was turned off. During each vocalization, participants heard their voice pitch randomly shifted +200 cents (100 cents = one semitone) while seeing a number of red indicator light flashes on the computer screen. The number of the pitch perturbations ranged from one to five per vocalization. The first pitch perturbation occurred 500–1000 ms after the onset of vocalization, and the succeeding stimuli were presented with an inter-stimulus ISI of 700–900 ms. The red indicator light flashed 1–13 times per vocalization. The first red indicator light began to flash 500 ms after the blue indicator light prompted participants to vocalize, and the succeeding stimuli were presented with three different ISIs: 1400–2000 ms (1400, 1600, 1800, and 2000 ms; low load), 900–1500 ms (900, 1100, 1300, and 1500 ms; intermediate load), and 400–1000 ms (400, 600, 800, and 1000 ms; high load). The onsets of auditory and visual stimuli were asynchronous. The durations of both the red indicator light and pitch perturbation were fixed at 200 ms. Production of  $\sim 40$  consecutive vocalizations constituted one block, which led to  $\sim 100$  trials (i.e., pitch perturbation) per condition. The order of the three attentional load conditions was counterbalanced across all subjects.

While producing sustained vocalizations, participants were required to divide their attention to auditory (pitch feedback perturbations) and visual stimuli (flashing red lights) across the three load conditions. An immediate recall test was performed after each vocalization, during which they reported the number of the pitch perturbations that they heard and the number of the red indicator light flashes that they saw. This test ensured that participants attended to the bimodal stimuli as required. Their behavioral performance, as indexed by the percentage of correctly remembered auditory and visual stimuli across the three load conditions, was evaluated and submitted to statistical analyses.

### Data Analysis

The event-related averaging technique (Li et al., 2013) was applied to the measurements of the magnitudes and latencies

of vocal responses to pitch perturbations using a custom-developed IGOR PRO software program (v.6.0 by Wavemetrics Inc.). First, the voice  $F_0$  contours in Hertz were extracted from the voice signals using Praat software (Boersma, 2001) and converted to the cents scale using the following formula: cents =  $100 \times (12 \times \log_2(F_0/\text{reference}))$  [reference = 195.997 Hz (G3)]. The voice contours in cents were then segmented into epochs of 200 ms before to 700 ms after the onset of the pitch perturbation. All individual trials were carefully inspected using a waterfall procedure and trials with signal processing errors or unexpected vocal stops were rejected from further analyses. Finally, the artifact-free trials were normalized by subtracting the mean  $F_0$  values in the baseline period (−200 to 0 ms) from the  $F_0$  values after the perturbation onset and then averaged to generate an overall response. The magnitude of a vocal response in cents was measured as the greatest  $F_0$  value following the response onset. The latency was defined as the time when the voice  $F_0$  contours exceeded 2 SDs above or below the pre-stimulus mean following the perturbation onset.

Cortical ERPs to pitch-shifted voice auditory feedback were measured using NetStation software. The EEG data were first band-pass filtered at 1–20 Hz and then segmented into epochs ranging from 200 ms before to 500 ms after the onset of the pitch perturbation. Following an artifact detection procedure, segmented trials with voltage values that exceeded  $\pm 55 \mu\text{V}$  of the moving average over an 80-ms window were rejected from further analysis. Additional visual inspection of all individual trials was performed to ensure that all trials with artifacts were removed. Individual electrodes were determined as bad electrodes if they contained artifacts in more than 20% of the segments, and any file that contained more than 10 bad electrodes was excluded. As a result, 88% of the trials were retained and re-referenced to the average of electrodes on each mastoid. Trials were then averaged and baseline-corrected to generate an overall ERP response for each condition. The amplitudes and latencies of N1 and P2 components (Hawco et al., 2009; Chen et al., 2012) were extracted as the negative and positive peaks in the time windows of 80–180 ms and 160–280 ms.

The magnitudes and latencies of vocal and cortical responses (N1 and P2) to pitch feedback perturbations were analyzed using repeated-measures analysis of variance (RM-ANOVAs) in SPSS (v. 16.0). The magnitudes and latencies of vocal responses were subjected to one-way RM-ANOVAs, in which attentional load (low, intermediate, and high load) was chosen as a within-subject factor. The amplitudes and latencies of N1 and P2 responses from 10 fronto-central electrodes (FC1, FC2, FCz, FC3, FC4, C1, C2, Cz, C3, and C4) were subjected to three-way RM-ANOVAs, including three within-subject factors of attentional load, anteriority, and laterality. Frontal (FC1, FC2, FCz, FC3, FC4) and central (C1, C2, Cz, C3, and C4) electrodes were chosen as an anteriority factor, while lateral left (FC3, C3), medial left (FC1, C1), midline (FCz, Cz), medial right (FC2, C2), and lateral right (FC4, C4) were used as a laterality factor. The Greenhouse-Geisser was used to correct probability values for multiple degrees of freedom when the assumption of sphericity was violated. Effect

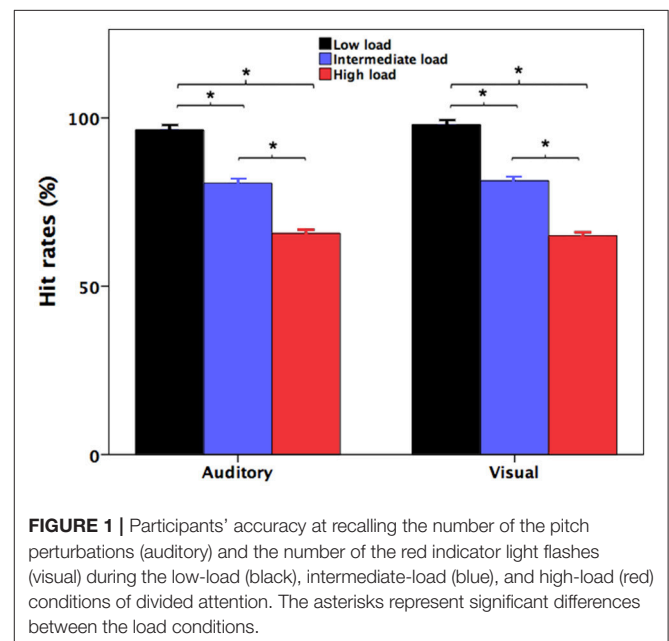
size was calculated using partial  $\eta^2$  to describe the size of differences between the conditions.  $P$ -values  $< 0.05$  and partial  $\eta^2 > 0.14$  (Richardson, 2011) were required to be considered significant.

## RESULTS

### Behavioral Performance

Figure 1 shows participants' accuracy at identifying the number of the pitch perturbations (auditory) and the number of the red indicator light flashes (visual) during divided attention as a function of attentional load. Participants' response accuracy for identifying the number of the red indicator light flashes in the high-load condition ( $65.0 \pm 1.0\%$ ; mean  $\pm$  standard errors of the mean throughout unless otherwise indicated) was significantly lower relative to both the intermediate-load ( $81 \pm 0.8\%$ ) [ $t_{(31)} = 12.852, p < 0.001$ ] and low-load conditions ( $98.0 \pm 0.4\%$ ) [ $t_{(31)} = 31.570, p < 0.001$ ]. Also, their response accuracy in the intermediate-load condition was also significantly lower than accuracy in the low-load condition [ $t_{(31)} = 19.909, p < 0.001$ ]. These results indicate that poorer behavioral performance was associated with faster presentation rate of the red indicator light.

Likewise, participants' accuracy for identifying the number of the pitch perturbations in the high-load condition ( $65.7 \pm 1.1\%$ ) was significantly lower than both the intermediate-load ( $80.6 \pm 0.9\%$ ) [ $t_{(31)} = 16.880, p < 0.001$ ] and low-load conditions ( $96.4 \pm 0.5\%$ ) [ $t_{(31)} = 25.790, p < 0.001$ ]. Their response accuracy in the intermediate-load condition was also significantly lower than that in the low-load condition [ $t_{(31)} = 10.725, p < 0.001$ ]. Therefore, response accuracy for identification of the number of the pitch perturbations was modulated by attentional load created by the different presentation rates of the red indicator lights that participants had to simultaneously count.



## Vocal Responses

**Figure 2A** shows the grand-averaged compensatory voice  $F_0$  contours in response to pitch perturbations across the three attentional loads. As can be seen, the high-load condition was associated with the largest vocal compensation, followed by the intermediate- and low-load conditions. A one-way RM-ANOVA conducted on the magnitudes of vocal responses revealed a significant main effect of attentional load [ $F_{(2, 62)} = 7.455, p = 0.004$ , partial  $\eta^2 = 0.194$ ]. *Post-hoc* Bonferroni comparison tests showed that the low-load condition ( $16.5 \pm 1.8$  cents) elicited significantly smaller response magnitudes than the intermediate-load ( $20.2 \pm 2.4$  cents) ( $p = 0.029$ ) and high-load conditions ( $23.5 \pm 4.0$  cents) ( $p = 0.015$ ) (see **Figure 2B**), while the difference between the intermediate-load and high-load conditions did not reach significance ( $p = 0.400$ ). In contrast, the latencies of vocal responses did not vary as a function of attentional load (low-load:  $134 \pm 13$  ms; intermediate-load:  $130 \pm 13$  ms; high-load:  $121 \pm 11$  ms) [ $F_{(2, 62)} = 0.371, p = 0.692$ , partial  $\eta^2 = 0.012$ ].

## ERP Responses

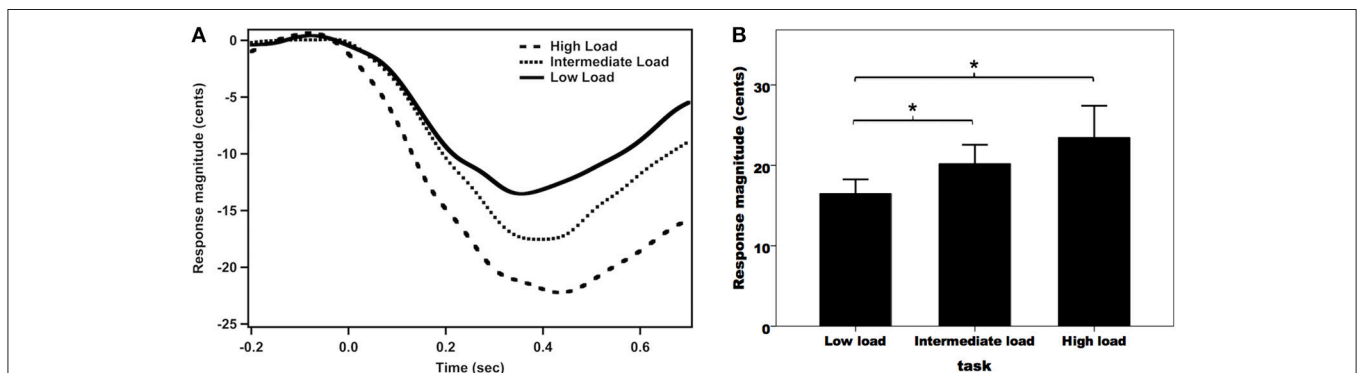
**Figure 3A** illustrates the grand-averaged ERP waveforms in response to pitch perturbations across the three attentional loads. Both the N1 and P2 response appeared to be affected by divided attention, as reflected by increased N1 responses and decreased P2 response with the increasing of attentional load. These effects of divided attention can also be seen in the topographical distributions of the N1 (**Figure 3B**) and P2 amplitudes (**Figure 3C**). A three-way RM-ANOVA conducted on the N1 amplitudes revealed a significant main effect of attentional load [ $F_{(2, 62)} = 8.744, p = 0.001$ , partial  $\eta^2 = 0.215$ ]. *Post-hoc* Bonferroni comparison tests showed significantly larger N1 amplitudes (more negative) in the high-load condition relative to the intermediate-load ( $p = 0.009$ ) and low-load conditions ( $p = 0.002$ ) (see **Figure 4A**), while N1 amplitudes in the low-load and intermediate-load conditions did not differ significantly ( $p = 1.000$ ). Larger N1 amplitudes for the frontal electrodes relative to the central electrodes led to a significant main effect of anteriority [ $F_{(1, 31)} = 16.550, p < 0.001$ , partial  $\eta^2 = 0.348$ ]. There was a

significant main effect of laterality [ $F_{(4, 124)} = 8.527, p < 0.001$ , partial  $\eta^2 = 0.216$ ], which was caused by smaller N1 amplitudes for the left medial electrodes relative to the left lateral ( $p < 0.001$ ) and central electrodes ( $p < 0.001$ ).

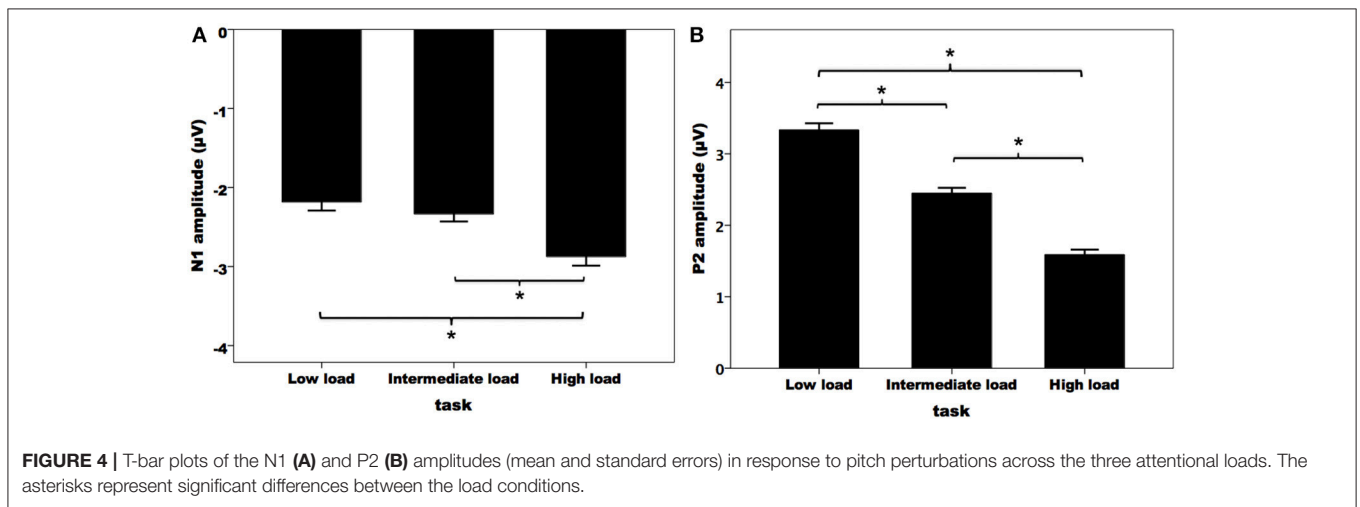
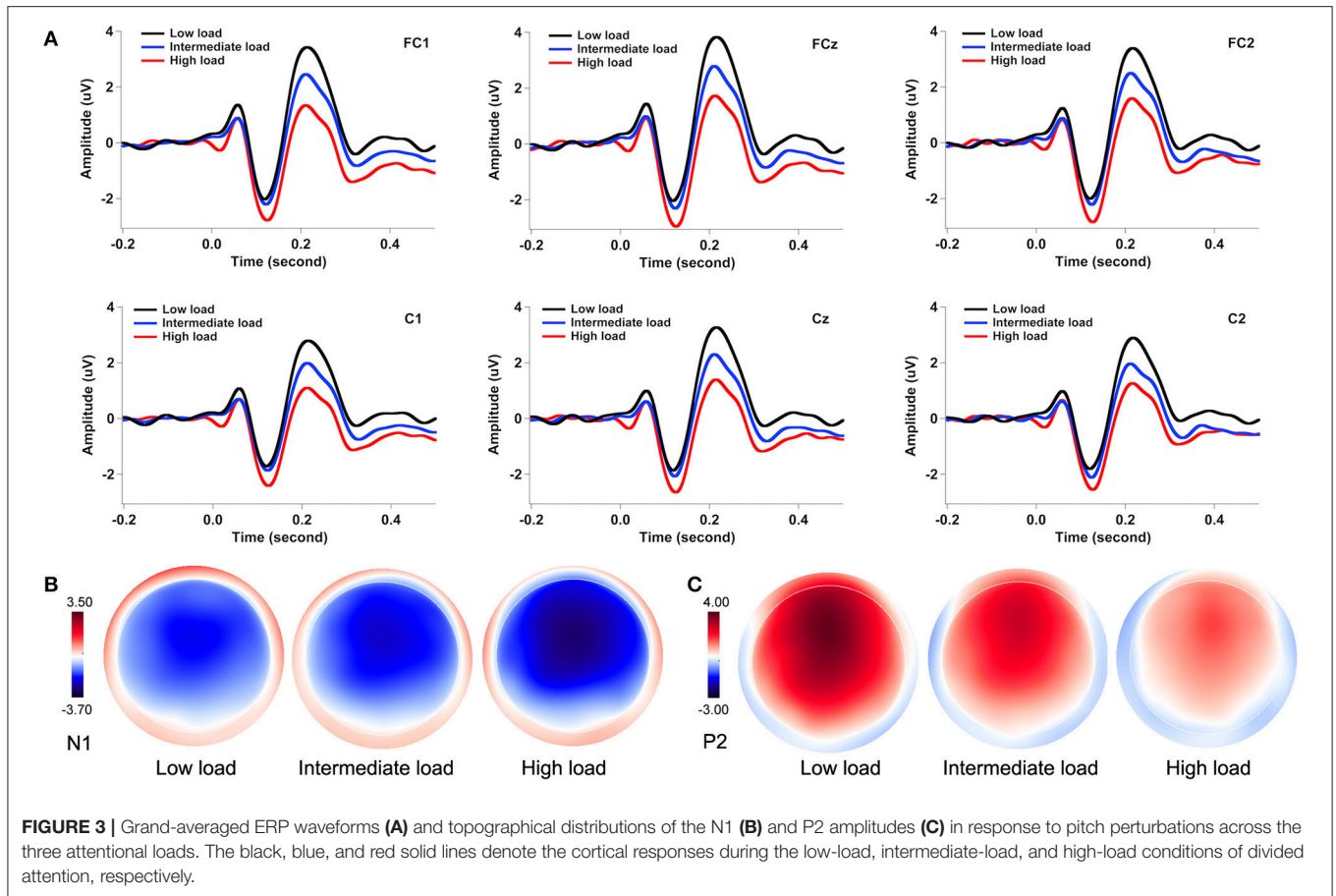
For the N1 latencies, the main effects of attentional load [ $F_{(2, 62)} = 1.426, p = 0.248$ , partial  $\eta^2 = 0.044$ ] and anteriority [ $F_{(1, 31)} = 2.738, p = 0.108$ , partial  $\eta^2 = 0.081$ ] did not reach significance, whereas a significant main effect of laterality was observed [ $F_{(4, 124)} = 5.756, p = 0.008$ , partial  $\eta^2 = 0.157$ ]. *Post-hoc* Bonferroni comparison tests showed significantly longer N1 latencies for the right lateral electrodes relative to the medial left ( $p = 0.013$ ), medial right ( $p = 0.009$ ), and middle electrodes ( $p = 0.001$ ).

A three-way RM-ANOVA conducted on the P2 amplitudes revealed a significant main effect of attentional load [ $F_{(2, 62)} = 91.495, p < 0.001$ , partial  $\eta^2 = 0.747$ ]. *Post-hoc* Bonferroni comparison tests showed that P2 amplitude was smaller in the high-load condition as compared to the intermediate-load condition ( $p < 0.001$ ) and low-load condition ( $p < 0.001$ ), and P2 amplitude was also smaller in the intermediate-load condition than the low-load condition ( $p < 0.001$ ) (see **Figure 4B**). The main effect of anteriority [ $F_{(1, 31)} = 46.923, p < 0.001$ , partial  $\eta^2 = 0.602$ ] reached significance, as reflected by significantly smaller P2 responses for the central electrodes relative to the frontal electrodes. There was also a significant main effect of laterality [ $F_{(4, 124)} = 27.066, p < 0.001$ , partial  $\eta^2 = 0.466$ ] that was the result of larger P2 amplitudes for the middle electrodes relative to the other electrodes ( $p < 0.02$ ) and larger P2 amplitudes for the medial electrodes relative to the lateral electrodes ( $p < 0.03$ ).

For the P2 latencies, there were no significant main effects of attentional load [ $F_{(2, 62)} = 1.805, p = 0.173$ , partial  $\eta^2 = 0.055$ ] and anteriority [ $F_{(1, 31)} = 0.200, p = 0.658$ , partial  $\eta^2 = 0.006$ ]. However, P2 latencies were modulated as a function of laterality [ $F_{(4, 124)} = 10.174, p < 0.001$ , partial  $\eta^2 = 0.247$ ], as reflected by significantly longer P2 latencies at the lateral right electrodes relative to than the medial left ( $p = 0.002$ ), lateral left ( $p = 0.012$ ), medial right ( $p = 0.003$ ), and middle electrodes ( $p = 0.001$ ).



**FIGURE 2 |** Grand-averaged voice  $F_0$  contours (**A**) and T-bar graphs of the absolute values of compensatory vocal responses (**B**) to pitch perturbations across the three attentional loads. The thick solid line, the dense dashed line, and the sparse dashed line represent the vocal responses during the low-load, intermediate-load, and high-load conditions of divided attention, respectively. The asterisks represent significant differences between the load conditions.



## DISCUSSION

By asking participants to attend to pitch perturbations in their voice auditory feedback while concurrently performing a low-load, intermediate-load, and high-load visual attention task, the present cross-modal study investigated the auditory-motor processing of vocal pitch errors during divided attention.

The behavioral results revealed significantly smaller vocal compensations for attended pitch perturbations in the low-load condition relative to the intermediate-load and high-load conditions. Differential effects of divided attention were observed on the cortical N1 and P2 responses to attended pitch perturbations. The high-load condition elicited significantly larger N1 responses and smaller P2 responses than the

intermediate-load and low-load conditions. These findings provide behavioral and neural evidence that divided attention can modulate the auditory-motor processing of vocal pitch errors.

In a previous study by Liu et al. (2015), we showed that dividing attention between pitch perturbations and flashing lights elicited significantly larger N1 responses and smaller P2 responses to pitch perturbations relative to passively observing the bimodal stimuli. In the present study, we found that both N1 and P2 responses to pitch perturbations were differentially modulated by divided attention, with larger N1 and smaller P2 responses elicited by higher attentional loads. These findings add further support to the idea that these two ERP components play different roles in the cortical processing of voice pitch regulation (Behroozmand et al., 2011; Hu et al., 2015; Guo et al., 2017). As important, increased load of divided attention elicited significantly enhanced vocal compensations for pitch perturbations. These findings provide the first behavioral evidence for the modulatory effects of divided attention on auditory feedback control of vocal production. Note that the vocal compensations between the intermediate-load and high-load conditions were not significantly different, nor were the differences of N1 amplitudes between the intermediate-load and low-load conditions significant. Nevertheless, the high-load condition elicited significantly larger vocal and N1 responses and smaller P2 responses than the low-load condition. Thus, the modulatory effect of divided attention on auditory-vocal integration appears to be subject to the degree of attentional load.

Given that attentional capacity is limited (Cowan et al., 2005), one might predict that increasing the presentation rate of the red indicator light flashes would produce increased demands on attention, which would in turn reduce the attentional resources available for identification of the number of pitch perturbations. The reduced attentional resources allocated to pitch feedback errors during the high-load vs. low-load condition should lead to decreased vocal compensations and cortical P2 responses, since focused attention elicits enhanced vocal and cortical P2 responses to pitch perturbations (Tumber et al., 2014; Hu et al., 2015; Liu et al., 2015). Paradoxically, however, the high-load condition elicited enhanced N1 responses and vocal compensations but suppressed P2 responses relative to the low-load condition. An important question comes from our findings then: what are the possible mechanisms underlying these differential neurobehavioral effects of divided attention on the auditory-motor processing of vocal pitch regulation?

One possible account is that these modulatory effects may reflect the interaction between working memory and divided attention in auditory feedback control of speech production. This interpretation is motivated by the fact that working memory is required to store and process multiple independent sensory stimuli during divided attention (Fagioli and Macaluso, 2009; Santangelo and Macaluso, 2013). The prefrontal cortex, which has been implicated in subserving working memory (Curtis and D'Esposito, 2004), is additionally recruited or more active during divided attention as compared to selective attention (Loose et al., 2003; Johnson and Zatorre, 2006; Moissala et al., 2015). Furthermore, brain regions that are involved in working memory are more active when load is increased during divided

attention tasks (Uncapher and Rugg, 2005; Santangelo and Macaluso, 2013; Oren et al., 2016). For example, Oren et al. (2016) asked participants to watch movies while simultaneously detecting whether a string of letters was a word or pseudo-word, during which attentional load was manipulated by making the lexical decision task easy and hard. As compared to the low-load condition, the high-load condition was associated with increased activation of the prefrontal cortex (Oren et al., 2016). In another study, Santangelo and Macaluso (2013) required participants to monitor both the object and location of the items. They found that increasing the load of divided attention led to a linear increase in brain activity in the intraparietal sulcus, a brain region that has been activated consistently in working memory studies (Majerus et al., 2007; Harrison et al., 2010). These findings suggest that divided attention and working memory may share a capacity-limited pool of neural resources (Santangelo and Macaluso, 2013). Returning to the present study, it is reasonable to hypothesize that increasing the presentation rate of the red indicator light flashes led to the allocation of more working memory resources to the online processing of both the red indicator light flashes and pitch feedback perturbations.

Along similar lines, recent evidence has shown the effects of working memory on auditory-motor integration for vocal pitch regulation. For example, Guo et al. (2017) reported that enhanced N1 responses in the left middle and superior temporal gyrus, and suppressed P2 responses in the left middle and superior temporal gyrus, inferior parietal lobule, somatosensory cortex, right inferior frontal gyrus and insula were elicited by a delayed match-to-sample (DMS) task that required participants to indicate whether the pitch perturbations they heard during vocalizations in test and sample sequences matched or not. And a significant positive correlation between improved working memory capacity and enhanced P2 responses was found for participants who underwent a training based on a digit-span backward (DSB) paradigm (Li et al., 2015). Considering that precise representations of auditory working memory information can be stored in the auditory cortex (Scott et al., 2014; Huang et al., 2016), enhanced N1 responses in the auditory regions reflect an allocation of more auditory working memory resources to the detection of mismatches between predicted and actual feedback during vocal production. Significant demands on working memory for the storage of pitch perturbations, however, reduce the availability of working memory resources for the auditory-motor transformations, as reflected by suppressed P2 responses in the fronto-parietal regions. In light of this account and the above overlapping hypothesis of divided attention and working memory, our findings of enhanced N1 responses and suppressed P2 responses with increasing attentional load may reflect the engagement of working memory in divided attention, suggestive of increased working memory resources available for the detection of pitch feedback errors but decreased working memory resources available for the auditory-motor transformations. This speculation is supported by one study by Uncapher and Rugg (2005) that required participants to judge whether the words on the screen represented a living or a nonliving thing while attending to an easy and hard auditory task. They found increased activity in the middle occipital cortex

and fusiform gyrus and decreased activity in the fronto-parietal regions during the hard vs. easy auditory task.

In addition to the modulation of cortical N1 and P2 responses, we also found enhanced vocal compensations for pitch perturbations in the intermediate-load and high-load conditions relative to the low-load condition. Interestingly, there was also a significant increase of vocal compensations for pitch perturbations in the DMS task that engaged working memory (Guo et al., 2017). Moreover, participants who received extensive auditory working memory training based on a frequency-pattern recognition (FPR) paradigm produced suppressed vocal compensations that were significantly correlated with improved working memory capacity, and enhanced P2 responses in the left middle frontal gyrus, inferior parietal lobule, right inferior frontal gyrus, and insula (Guo et al., 2017). These regions are not only involved in working memory but also in inhibitory control (Aron et al., 2004; Barber et al., 2013; Chmielewski et al., 2017), an important cognitive function that depends on the amount of working memory resources to inhibit reflex-like behavioral responses (Barber et al., 2013; Chmielewski et al., 2015). It is thus suggested that working memory can inhibit compensatory vocal adjustment to prevent vocal production from being excessively influenced by auditory feedback (Guo et al., 2017). In light of these findings, enhanced vocal compensations with increasing attentional load observed in the present study can be accounted for as a result of impaired inhibitory control processes caused by reduced working memory resources for the auditory-motor transformations as reflected by suppressed P2 responses.

It should be noted, however, that our interpretation of the interaction between divided attention and working memory in auditory-vocal integration is speculative. For example, working memory was not directly measured or specifically manipulated in the present study. In addition, whether the observed changes in the N1 and P2 responses to pitch perturbations across the attentional loads received contributions from the neural substrates involved in auditory working memory is unknown due to lack of knowledge about the neural generators of these two ERP components.

Future neuroimaging experiments, where participants divide attention to different sensory stimuli while maintaining their specific features (e.g., category, location, etc.) in working memory, should be conducted to verify our speculation.

In summary, the present cross-modal study investigated the behavioral and neural correlates of auditory-motor integration for vocal pitch regulation during divided attention. The results revealed enhanced vocal compensations for pitch perturbations, enhanced N1 responses, and suppressed P2 responses with increasing load of divided attention, providing neurobehavioral evidence that divided attention can exert top-down influences on auditory feedback control of speech production. Considering the involvement of working memory in divided attention for the storage and maintenance of multiple sensory information (Johnson and Zatorre, 2006; Johnson et al., 2007; Santangelo and Macaluso, 2013), our findings may reflect the contribution of working memory to auditory-vocal integration during divided attention.

## AUTHOR CONTRIBUTIONS

HL and BZ: Designed the experiment; YL, HF, JL, and PL: Performed the experiment and analyzed the data; YL, JJ, BZ, and HL: Interpreted the results and wrote the manuscript. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

This study was funded by grants from the National Natural Science Foundation of China (Nos. 31371135, 81472154, and 81772439), Guangdong Natural Science Funds for Distinguished Young Scholar (No. S2013050014470), Guangdong Province Science and Technology Planning Project (No. 2017A050501014), Guangzhou Science, and Technology Programme (No. 201604020115), and the Fundamental Research Funds for the Central Universities (No. 15ykjc13b). YL, HF, and JL contributed equally to this work.

## REFERENCES

- Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14608–14613. doi: 10.1073/pnas.0510480103
- Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* 8, 170–177. doi: 10.1016/j.tics.2004.02.010
- Barber, A. D., Caffo, B. S., Pekar, J. J., and Mostofsky, S. H. (2013). Effects of working memory demand on neural mechanisms of motor response selection and control. *J. Cogn. Neurosci.* 25, 1235–1248. doi: 10.1162/jocn\_a\_00394
- 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
- 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., Liu, H., and Larson, C. R. (2011). Time-dependent neural processing of auditory feedback during voice pitch error detection. *J. Cogn. Neurosci.* 23, 1205–1217. doi: 10.1162/jocn.2010.21447
- 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
- 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., 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
- Chmielewski, W. X., Mückschel, M., Stock, A. K., and Beste, C. (2015). The impact of mental workload on inhibitory control subprocesses. *Neuroimage* 112, 96–104. doi: 10.1016/j.neuroimage.2015.02.060
- Chmielewski, W. X., Mückschel, M., Ziemssen, T., and Beste, C. (2017). The norepinephrine system affects specific neurophysiological subprocesses in the modulation of inhibitory control by working memory demands. *Hum. Brain Mapp.* 38, 68–81. doi: 10.1002/hbm.23344
- Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51, 42–100. doi: 10.1016/j.cogpsych.2004.12.001
- Craik, F. I., Govoni, R., Naveh-Benjamin, M., and Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *J. Exp. Psychol. Gen.* 125, 159–180. doi: 10.1037/0096-3445.125.2.159
- Curtis, C. E., and D'Esposito, M. (2004). The effects of prefrontal lesions on working memory performance and theory. *Cogn. Affect. Behav. Neurosci.* 4, 528–539. doi: 10.3758/CABN.4.4.528
- Fagioli, S., and Macaluso, E. (2009). Attending to multiple visual streams: interactions between location-based and category-based attentional selection. *J. Cogn. Neurosci.* 21, 1628–1641. doi: 10.1162/jocn.2009.21116
- Ferree, T. C., Luu, P., Russell, G. S., and Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clin. Neurophysiol.* 112, 536–544. doi: 10.1016/S1388-2457(00)00533-2
- Getzmann, S., Golob, E. J., and Wascher, E. (2016). Focused and divided attention in a simulated cocktail-party situation: ERP evidence from younger and older adults. *Neurobiol. Aging* 41, 138–149. doi: 10.1016/j.neurobiolaging.2016.02.018
- 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
- Guo, Z., Wu, X., Li, W., Jones, J. A., Yan, N., Sheft, S., et al. (2017). Top-down modulation of auditory-motor integration during speech production: the role of working memory. *J. Neurosci.* 37, 10323–10333. doi: 10.1523/JNEUROSCI.1329-17.2017
- Harrison, A., Jolicoeur, P., and Marois, R. (2010). “What” and “where” in the intraparietal sulcus: an fMRI study of object identity and location in visual short-term memory. *Cereb. Cortex* 20, 2478–2485. doi: 10.1093/cercor/bhp314
- Hawco, C. S., Jones, J. A., Ferretti, T. R., and Keough, D. (2009). ERP correlates of online monitoring of auditory feedback during vocalization. *Psychophysiology* 46, 1216–1225. doi: 10.1111/j.1469-8986.2009.00875.x
- Hickok, G., Houde, J. F., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422. doi: 10.1016/j.neuron.2011.01.019
- Hillyard, S. A., Vogel, E. K., and Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond B. Biol. Sci.* 353, 1257–1270. doi: 10.1098/rstb.1998.0281
- Hink, R. F., and Hillyard, S. A. (1976). Auditory evoked potentials during selective listening to dichotic speech messages. *Percept. Psychophys.* 20, 236–242. doi: 10.3758/BF03199449
- Hu, H., Liu, Y., Guo, Z., Li, W., Liu, P., Chen, S., et al. (2015). Attention modulates cortical processing of pitch feedback errors in voice control. *Sci. Rep.* 5:7812. doi: 10.1038/srep07812
- Huang, Y., Matysiak, A., Heil, P., König, R., and Brosch, M. (2016). Persistent neural activity in auditory cortex is related to auditory working memory in humans and nonhuman primates. *Elife* 5:e15441. doi: 10.7554/eLife.15441
- Johnson, J. A., Strafella, A. P., and Zatorre, R. J. (2007). The role of the dorsolateral prefrontal cortex in bimodal divided attention: two transcranial magnetic stimulation studies. *J. Cogn. Neurosci.* 19, 907–920. doi: 10.1162/jocn.2007.19.6.907
- Johnson, J. A., and Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage* 31, 1673–1681. doi: 10.1016/j.neuroimage.2006.02.026
- Jones, J. A., and Munhall, K. G. (2002). The role of auditory feedback during phonation: studies of Mandarin tone production. *J. Phon.* 30, 303–320. doi: 10.1006/jpho.2001.0160
- 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
- Klingberg, T. (1998). Concurrent performance of two working memory tasks: potential mechanisms of interference. *Cereb. Cortex* 8, 593–601. doi: 10.1093/cercor/8.7.593
- Larson, C. R. (1998). Cross-modality influences in speech motor control: the use of pitch shifting for the study of F0 control. *J. Commun. Disord.* 31, 489–503. doi: 10.1016/S0021-9924(98)00021-5
- 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
- Li, W., Guo, Z., Jones, J. A., Huang, X., Chen, X., Liu, P., et al. (2015). Training of working memory impacts neural processing of vocal pitch regulation. *Sci. Rep.* 5:16562. doi: 10.1038/srep16562
- 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., 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
- Liu, H., Wang, E. Q., Verhagen Metman, L., and Larson, C. R. (2012). Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson's disease. *PLoS ONE* 7:e33629. doi: 10.1371/journal.pone.0033629
- Liu, H., Zhang, Q., Xu, Y., and Larson, C. R. (2007). Compensatory responses to loudness-shifted voice feedback during production of Mandarin speech. *J. Acoust. Soc. Am.* 122, 2405–2412. doi: 10.1121/1.2773955
- Liu, Y., Hu, H., Jones, J. A., Guo, Z., Li, W., Chen, X., et al. (2015). Selective and divided attention modulates auditory-vocal integration in the processing of pitch feedback errors. *Eur. J. Neurosci.* 42, 1895–1904. doi: 10.1111/ejn.12949
- Loose, R., Kaufmann, C., Auer, D. P., and Lange, K. W. (2003). Human prefrontal and sensory cortical activity during divided attention tasks. *Hum. Brain Mapp.* 18, 249–259. doi: 10.1002/hbm.10082
- Macdonald, E. N., Goldberg, R., and Munhall, K. G. (2010). Compensations in response to real-time formant perturbations of different magnitudes. *J. Acoust. Soc. Am.* 127, 1059–1068. doi: 10.1121/1.3278606
- Majerus, S., Bastin, C., Poncelet, M., Van der Linden, M., Salmon, E., Collette, F., et al. (2007). Short-term memory and the left intraparietal sulcus: focus of attention? Further evidence from a face short-term memory paradigm. *Neuroimage* 35, 353–367. doi: 10.1016/j.neuroimage.2006.12.008
- Mitsuya, T., MacDonald, E. N., Munhall, K. G., and Purcell, D. W. (2015). Formant compensation for auditory feedback with English vowels. *J. Acoust. Soc. Am.* 138, 413. doi: 10.1121/1.4923154
- Mitsuya, T., Samson, F., Ménard, L., and Munhall, K. G. (2013). Language dependent vowel representation in speech production. *J. Acoust. Soc. Am.* 133, 2993–3003. doi: 10.1121/1.4795786
- Moisala, M., Salmela, V., Salo, E., Carlson, S., Vuontela, V., Salonen, O., et al. (2015). Brain activity during divided and selective attention to auditory and visual sentence comprehension tasks. *Front. Hum. Neurosci.* 9:86. doi: 10.3389/fnhum.2015.00086
- Mollaei, F., Shiller, D. M., Baum, S. R., and Gracco, V. L. (2016). Sensorimotor control of vocal pitch and formant frequencies in Parkinson's disease. *Brain Res.* 1646, 269–277. doi: 10.1016/j.brainres.2016.06.013
- Möttönen, R., van de Ven, G. M., and Watkins, K. E. (2014). Attention fine-tunes auditory-motor processing of speech sounds. *J. Neurosci.* 34, 4064–4069. doi: 10.1523/JNEUROSCI.2214-13.2014
- 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

- 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
- Naveh-Benjamin, M., Craik, F. I., Gavrilescu, D., and Anderson, N. D. (2000). Asymmetry between encoding and retrieval processes: evidence from divided attention and a calibration analysis. *Mem. Cognit.* 28, 965–976. doi: 10.3758/BF03209344
- Oren, N., Shapira-Lichter, I., Lerner, Y., Tarrasch, R., Hendler, T., Giladi, N., et al. (2016). How attention modulates encoding of dynamic stimuli. *Front. Hum. Neurosci.* 10:507. doi: 10.3389/fnhum.2016.00507
- 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., Mizuiri, 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
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educ. Res. Rev.* 6, 135–147. doi: 10.1016/j.edurev.2010.12.001
- Sabri, M., Binder, J. R., Desai, R., Medler, D. A., Leitzl, M. D., and Liebenthal, E. (2008). Attentional and linguistic interactions in speech perception. *Neuroimage* 39, 1444–1456. doi: 10.1016/j.neuroimage.2007.09.052
- Santangelo, V., and Macaluso, E. (2013). The contribution of working memory to divided attention. *Hum. Brain Mapp.* 34, 158–175. doi: 10.1002/hbm.21430
- Scott, B. H., Mishkin, M., and Yin, P. (2014). Neural correlates of auditory short-term memory in rostral superior temporal cortex. *Curr. Biol.* 24, 2767–2775. doi: 10.1016/j.cub.2014.10.004
- Stevens, C., Sanders, L., and Neville, H. (2006). Neurophysiological evidence for selective auditory attention deficits in children with specific language impairment. *Brain Res.* 1111, 143–152. doi: 10.1016/j.brainres.2006.06.114
- Tumber, A. K., Scheerer, N. E., and Jones, J. A. (2014). Attentional demands influence vocal compensations to pitch errors heard in auditory feedback. *PLoS ONE* 9:e109968. doi: 10.1371/journal.pone.0109968
- Uncapher, M. R., and Rugg, M. D. (2005). Effects of divided attention on fMRI correlates of memory encoding. *J. Cogn. Neurosci.* 17, 1923–1935. doi: 10.1162/089892905775008616
- 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

**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 © 2018 Liu, Fan, Li, Jones, Liu, Zhang 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 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.