



## OPEN

## Attention Modulates Cortical Processing of Pitch Feedback Errors in Voice Control

SUBJECT AREAS:  
SENSORY PROCESSING  
ATTENTIONHuijing Hu<sup>1,2\*</sup>, Ying Liu<sup>1\*</sup>, Zhiqiang Guo<sup>3</sup>, Weifeng Li<sup>1</sup>, Peng Liu<sup>1</sup>, Shaozhen Chen<sup>1</sup> & Hanjun Liu<sup>1,3</sup><sup>1</sup>Department of Rehabilitation Medicine The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China 510080, <sup>2</sup>Guangdong Provincial Work Injury Rehabilitation Center, Guangzhou, China 510440, <sup>3</sup>Department of Biomedical Engineering, School of Engineering, Sun Yat-sen University, Guangzhou, China 510006.Received  
3 September 2014Accepted  
10 December 2014Published  
15 January 2015

Correspondence and requests for materials should be addressed to H.J.L. (lhjun@mail.sysu.edu.cn)

\* These authors contributed equally to this work.

Considerable evidence has shown that unexpected alterations in auditory feedback elicit fast compensatory adjustments in vocal production. Although generally thought to be involuntary in nature, whether these adjustments can be influenced by attention remains unknown. The present event-related potential (ERP) study aimed to examine whether neurobehavioral processing of auditory-vocal integration can be affected by attention. While sustaining a vowel phonation and hearing pitch-shifted feedback, participants were required to either ignore the pitch perturbations, or attend to them with low (counting the number of perturbations) or high attentional load (counting the type of perturbations). Behavioral results revealed no systematic change of vocal response to pitch perturbations irrespective of whether they were attended or not. At the level of cortex, there was an enhancement of P2 response to attended pitch perturbations in the low-load condition as compared to when they were ignored. In the high-load condition, however, P2 response did not differ from that in the ignored condition. These findings provide the first neurophysiological evidence that auditory-motor integration in voice control can be modulated as a function of attention at the level of cortex. Furthermore, this modulatory effect does not lead to a general enhancement but is subject to attentional load.

The integration of sensory information arising from different pathways such as auditory, visual, and somatosensory system into the vocal motor system is critical for the control of speech production. Speech motor control is heavily dependent on the input from the central auditory system, i.e., auditory feedback<sup>1,2</sup>. In the past two decades, the altered auditory feedback (AAF) paradigm has been developed and widely used to investigate the mechanisms underlying auditory feedback control of vocal production<sup>3,4</sup>. Numerous studies have demonstrated that people generally produce compensatory vocal changes in the direction opposite to that of perceived pitch/loudness/formant perturbations in voice auditory feedback, although a few responses change in the same direction as the perturbation<sup>5–8</sup>. Multiple lines of evidence have shown that this feedback-based mechanism can be modulated as a function of auditory stimuli<sup>9,10</sup>, experimental task<sup>8,11</sup>, and learning experience<sup>12,13</sup>. Despite this, mechanisms underlying auditory feedback control of voice remain poorly understood.

Typically, vocal compensation for auditory feedback perturbations is exhibited as a corrective response with a short latency approximately ranging from 80–150 ms<sup>5,6</sup>. Behavioral studies have shown that vocally-untrained participants produce vocal compensation for pitch feedback perturbations even when told to ignore them throughout the experiment<sup>14,15</sup>, and they are unable to suppress vocal responses to pitch feedback perturbations irrespective of being told to ignore or compensate for them<sup>13,16</sup>. Therefore, auditory-motor integration in voice control is generally thought to be involuntary in nature and does not appear to be modified by cognitive function such as attention.

According to the internal forward model theory, however, the mechanisms underlying auditory feedback control of voice is a top-down controlled process<sup>17–19</sup>. In this model, a copy of motor command (i.e. efference copy) is generated to predict the sensory consequences of self-produced vocalization via auditory feedback, which is then compared with the actual auditory feedback (i.e. re-efference) to monitor the state of vocal production. Once there is a mismatch between the intended and actual auditory feedback, a new motor command is generated to correct for perturbations in auditory feedback to stabilize vocal production. Several neurophysiological studies have provided supportive evidence for this hypothesis that neural activity in the auditory cortex in response to pitch perturbations during active vocalization is enhanced relative to that during passive listening<sup>18,20,21</sup>, and the activity of brain regions that produce enhanced responses to pitch perturbation is predictive of vocal compensation behavior<sup>22</sup>.



Moreover, detection of perturbations in voice auditory feedback at the level of cortex is key to generate the corrective motor command in stabilizing vocal production. Although perception of speech sounds is considered as a highly automatic process<sup>23,24</sup>, considerable evidence has demonstrated the attention-driven auditory processing of speech sounds. For example, attending to an auditory stimulus leads to greater event-related potentials (ERPs) of N1<sup>25–28</sup> and P2<sup>26,29,30</sup> and enhanced brain activity<sup>31,32</sup> in the auditory cortex as compared to when ignoring it. Moreover, this enhancement effect of attention is suggested to be “elastic” and occurs as a function of attentional load<sup>33</sup>. For example, when comparing to the ignored condition, ERP amplitudes of P170 to attended stimuli were significantly increased during the low-load condition but declined during the high-load condition<sup>34</sup>. Recently, using transcranial magnetic stimulation (TMS) and magnetoencephalography (MEG) techniques, Mottonen et al.<sup>35</sup> reported that TMS-induced disruption of the lip representation increased left-hemisphere P50m response to attended sounds, providing supportive evidence that attention can modulate the auditory-motor interaction during speech processing. In light of these findings, focused auditory attention may lead to an increase of brain activity in the auditory-motor processing of feedback information, leaving open the possibility that there may be attention-driven mechanisms underlying auditory feedback control of vocal production at the level of cortex. This question, however, has yet to be answered.

In addition, several neuroimaging and neuromagnetic studies have been conducted to investigate the neural substrates involved in the auditory-feedback-based voice control<sup>13,36–38</sup>. Results of these studies identified brain regions including superior temporal gyrus (STG), posterior superior temporal sulcus (pSTS), anterior cingulate cortex (ACC), dorsal premotor cortex (dPMC), left supermarginal gyrus, inferior frontal gyrus (IFG) and anterior insula. Through the recording of electrocorticogram (ECoG) data from epilepsy patients, Chang et al.<sup>22</sup> and Greenlee et al.<sup>39</sup> found that pitch perturbations elicited greater brain activity in STG during vocalization as compared to listening. On the other hand, a cortical network of attention resources involved in the auditory processing of sounds includes STS, STG, inferior parietal lobules, IFG and premotor/supplementary motor cortices<sup>40–42</sup>. As can be seen, some brain regions such as STS, STG and IFG are involved in both auditory attention and vocal motor control, indicating that there is an overlapping of neural networks between them. Through this shared framework, it is plausible to suggest that there may be a modulatory effect of attention on the auditory-motor processing of sounds during vocal production.

Therefore, the present ERP study aimed to investigate whether auditory-motor integration in voice control can be modulated by auditory attention at the levels of both behavior and cortex. The AAF paradigm was used to perturb participants’ voice pitch feedback when they sustained a vowel phonation, during which they were asked to attend to or ignore those perturbations. In order to testify if effects of attention on the processing of pitch perturbations in vocal motor control are subject to task demands, the load of auditory attention in the attended conditions was manipulated in the present study (i.e. low vs. high level). Vocal and neurophysiological responses to pitch perturbations were measured across different attention tasks. Given that ERP components of N1 and P2 are reliably elicited by pitch perturbations in voice auditory feedback and reflect the auditory-motor integration of voice control at the cortical level<sup>21,43,44</sup>, they were extracted in the present study to index change of cortical response as a function of attention. Considering the reflexive-like nature of vocal response to pitch perturbations, we expected no systematic change of vocal response as a function of attention. We also expected that attending to pitch perturbations would result in increased amplitudes of N1 and P2 as compared to when they were ignored. Moreover, since attention effect on the central auditory processing is subject to task demands, we predicted that N1 and P2 responses would vary as a function of attentional load.

## Methods

**Subjects.** Twenty-eight female young adults, who were students from Sun Yat-sen University of China, participated in the experiment. Seven of 28 subjects were excluded from the final sample due to following reasons: failure of performing the experiment as required (N = 2), musical training experience (N = 1), native-Cantonese speaker (N = 1), malfunction of the equipment (N = 1), and excessive artifacts of the EEG data (N = 2). Therefore, 21 subjects entered the final dataset. They were right handed, native-Mandarin speakers aged 20–27 years of old (mean = 21; standard deviation [SD] = 2). None of them reported speech, hearing, or neurological disorders. All subjects passed a hearing screening at the threshold of 25 dB hearing level (HL) for pure-tone frequencies of 500, 1000, 2000 and 4000 Hz, and reported normal or corrected-normal vision. Informed consent was obtained from all subjects in compliance with a research protocol approved by the Institute Review Board of The Affiliated Hospital at Sun Yat-sen University of China. The study was carried out in accordance with the approved guidelines.

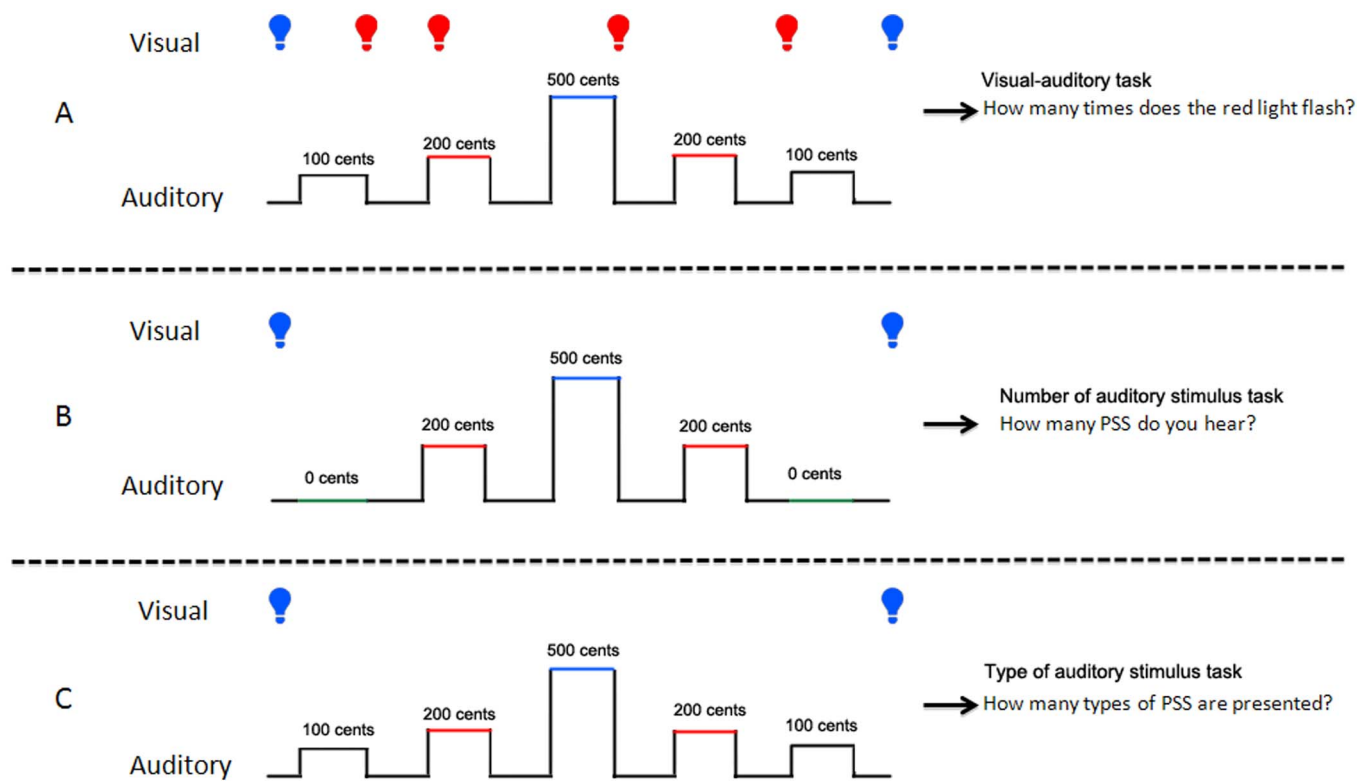
**Apparatus.** All subjects performed the experimental task in a sound-attenuated booth. A dynamic microphone (model DM2200, Takstar Inc.) was used to record the voice signals, which were then amplified with a MOTU Ultralite Mk3 firewire audio interface and an ICON NeoAmp headphone amplifier. The amplified signals were pitch-shifted by an Eventide Eclipse Harmonizer, and finally fed back to subjects through insert earphones (ER1-14A, Etymotic Research Inc.). The intensity of voice feedback had a gain of 10 dB sound pressure level (SPL) relative to that of vocal output to partially mask the air-borne and bone-conducted feedback. Max/MSP (v.5.0 by Cycling 74) software was used to control the Harmonizer to pitch-shift the voice feedback, in which acoustic parameters of the pitch-shift stimulus (PSS) including direction, magnitude, duration, and inter-stimulus intervals were manipulated. Transistor-transistor logic (TTL) pulses were generated by Max/MSP to mark the onset and offset of pitch perturbations. The original voice, pitch-shifted feedback, and TTL pulses were digitized at a sampling frequency of 10 kHz by a PowerLab A/D converter (model ML880, AD Instruments), and recorded using LabChart software (v.7.0 by AD Instruments). A custom-developed LED indicator light box was used to provide visual cues for controlling the onset and offset of vocalization and presenting the visual stimuli for the ignored condition (see below).

**Procedure.** Across all experimental conditions, subjects were instructed to vocalize a vowel sound/u/ following a visual cue (i.e. one blue indicator light). That is, they started vocalizing when the blue light was on and sustained a stable vocalization until the blue light was off. At the end of each vocalization, subjects were required to take a break of 2–3 s prior to initiating the next vocalization. The present study consisted of three tasks: visual-auditory (VA) task, number of auditory stimulus (NAS) task and type of auditory stimulus (TAS) task (see Figure 1). Across three tasks, subjects were given two practice blocks prior to the formal test. In the VA task, one red indicator light started flashing as long as subjects started vocalizing, during which their voice feedback was randomly pitch-shifted. The onsets of visual and auditory stimuli were asynchronous. The red light flashed up to 8 times with an inter-stimulus interval of 500–2000 ms (ISI). During each vocalization, voice feedback was pitch-shifted five times and the magnitude of PSS randomized from +100, +200, to +500 cents (100 cents = 1 semitone). The first PSS occurred 500–1000 ms after the vocal onset, and the succeeding stimuli were presented with an ISI of 700–900 ms. Totally, 60% of stimuli were for +200 cents and 20% for +100 and +500 cents, respectively. Subjects were required to count the number of the red light flashing per vocalization. The VA task serves as one ignored condition during which subjects ignored the auditory stimuli while attending to the visual stimuli.

In the following two tasks, the red indicator light was turned off ensuring that subjects would fully attend to the auditory stimuli. Load was manipulated by asking subjects to detect the specific features of the PSS. In the low-load condition (i.e. the NAS task), subjects were asked to count how many times voice feedback was pitch-shifted per vocalization. Voice pitch feedback was shifted five times per vocalization. Stimuli of 0 cent were mixed with stimuli of +200 and +500 cents, leading to a variable number of PSS ranging from 0 (i.e. all PSS were 0 cent) to 5 (i.e. none of PSS was 0 cent). Similar to the VA task, 60% of stimuli were for +200 cents and 20% for 0 and +500 cents.

In the high-load condition (i.e. the TAS task), subjects also sustained vocalization while hearing five pitch perturbations with three randomized magnitudes (+100, +200, and +500 cents). In contrast with the NAS task, subjects were required to count how many types (i.e. magnitude) of PSS they heard per vocalization. So subjects would hear 1 (i.e. only one magnitude was used) to 3 (i.e. all three magnitudes were used) types of stimuli per vocalization. Totally, 60% of them were for +200 cents and 20% for +100 and +500 cents respectively.

Note that stimulus of +100 cents was not used in the NAS task because pilot data showed that subjects were confused about +100 and +200 cents quite often, resulting in a difficulty in manipulating the load level between the NAS and TAS task. Previous research has shown that cortical responses (i.e. N1, P2) become greater as the magnitudes of pitch perturbation increase<sup>21,45</sup>. If responses to all three stimuli were averaged and used in the statistical analyses, it would be difficult to determine whether the change in cortical response was caused by attention manipulation or salient stimulus. In the present study, therefore, we examined the behavioral and neurophysiological responses to +200 cents only. For each task, subjects received 5 PSS over 35 consecutive vocalizations and 60% of them were for +200 cents, resulting in a total of 105 trials of +200 cents stimuli. The order of three tasks was randomized across all subjects. In addition, an immediate recalling test was performed after each



**Figure 1 | Schematic depicting the experimental setup.** Across all conditions, subjects were instructed to start and end vocalizing when the blue light was on and off. An immediate recalling test was performed after each vocalization for all subjects ensuring that the stimuli were attended or ignored as required. While sustaining a stable vocalization and hearing their voice unaltered (i.e. 0 cent) or pitch-shifted +100, +200, or +500 cents through insert earphones, subjects were asked to count how many times the red light flashed on the screen in the visual-auditory (VA) task (A), how many times their voice was pitch-shifted in the number of auditory stimulus (NAS) task (B), or how many types of PSS were presented in the type of auditory stimulus (TAS) task (C). Voice pitch feedback was shifted five times per vocalization. In the VA task, the red light flashed 1 to 8 times per vocalization. In the NAS task, stimuli of 0 cent were mixed with stimuli of +200 and +500 cents, leading to a variable number of PSS ranging from 0 (i.e. all PSS were 0 cent) to 5 (i.e. none of PSS was 0 cent) per vocalization. In the TAS task, the number of stimulus type ranged from 1 (i.e. only one magnitude was used) to 3 (i.e. all three magnitudes were used) per vocalization.

vocalization across all tasks for all subjects ensuring that the stimuli were attended to or ignored as required.

**Vocal response analyses.** Vocal responses to PSS were measured offline in one Macintosh computer. Voice fundamental frequency ( $F_0$ ) contours in Hz were extracted from the voice signals in Praat<sup>46</sup>, and converted to the cent scale in IGOR PRO (v.6.0, Wavemetrics Inc.) using the formula: cents =  $100 \times (39.86 \times \log_{10}(F_0/\text{reference}))$ . The reference is an arbitrary reference note of 195.997 Hz (G4). Voice  $F_0$  contours were then segmented into epochs ranging from 200 ms before and 700 ms after the onset of pitch perturbation. A visual inspection of all individual trials was used to reject those bad trials containing vocal interruption or signal processing errors. Artifact-free trials were averaged to generate an overall response for each stimulus condition. Measurement of vocal response was performed using the event-related averaging technique along with the pre-sorting method<sup>47,48</sup>. The response magnitude was calculated by subtracting the peak value of voice contour following the response onset from the pre-stimulus mean (−200 to 0 ms). The response latency was determined as the time of voice  $F_0$  departure from the pre-stimulus mean by more than 2 standard deviations (SDs).

**EEG data acquisition and analyses.** Subjects wore a 64-electrode Geodesic Sensor Net on the scalp throughout the recording. The data were amplified and digitized at a sampling frequency of 1 k Hz by Net Amps 300 amplifier (Electrical Geodesics Inc., Eugene, OR). The EEG signals were referenced to the vertex (Cz) for each channel during the online recording. Individual sensors were adjusted to maintain their impedances below 50 k $\Omega$  during the experiment<sup>49</sup>.

NetStation software (v.4.5, Electrical Geodesics Inc., Eugene, OR) was used to offline analyze the EEG signals across all conditions. All channels were re-referenced to the average of electrodes on each mastoid and band-pass filtered using a finite impulse filter (FIR) with cut-off frequencies set to 1–20 Hz (passband gain: 0.1 dB; stopband gain: 40 dB; rolloff frequency: 2 Hz). The continuous EEG was segmented into epochs with a window of −200 ms and +500 ms relative to the PSS onset, and then submitted to Artifact Detection toolbox in NetStation to reject those artifacts caused by excessive muscular activity, eye blinks, or eye movement. An additional visual inspection was performed to ensure the appropriate artifact rejection. Finally,

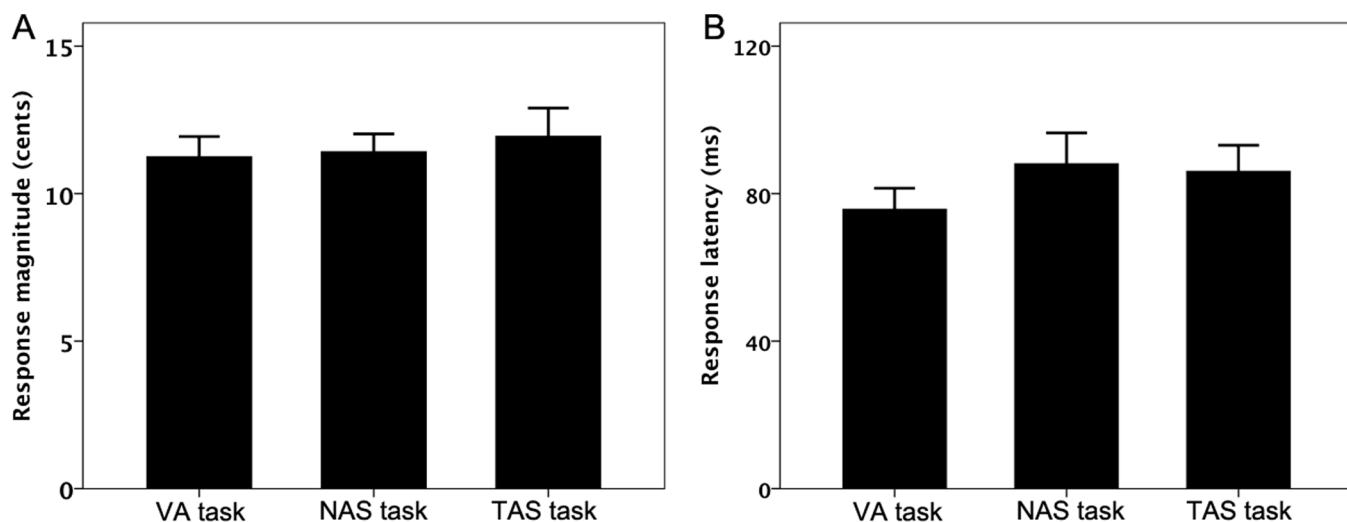
artifact-free trials were averaged and baseline-corrected to generate an overall response for each condition. The amplitudes and latencies of N1 and P2 components were measured as the negative and positive peaks in the time windows of 80–180 ms and 160–280 ms after the onset of PSS, respectively.

**Statistical analyses.** Values of vocal and neurophysiological responses to PSS across conditions were subjected to repeated-measures analyses of variance (RM-ANOVAs) using SPSS (v.16.0). The magnitudes and latencies of vocal response were analyzed using one-way RM-ANOVAs with the task (VA, NAS, TAS) as a within-subject factor. The amplitude and the latency of the N1-P2 complex, extracted from 15 electrodes (FC1, FC2, FCz, FC3, FC4, C1, C2, Cz, C3, C4, P1, P2, Pz, P3, P4), were subjected to three-way RM-ANOVAs (task, anteriority, and laterality). Frontal (FC1, FC2, FCz, FC3, FC4), central (C1, C2, Cz, C3, C4), and parietal (P1, P2, Pz, P3, P4) electrodes were chosen as an anteriority factor, whereas lateral left (FC3, C3, P3), medial left (FC1, C1, P1), midline (FCz, Cz, Pz), medial right (FC2, C2, P2), lateral right (FC4, C4, P4) were used as a laterality factor. Violation of the assumption of sphericity would lead to the correction of probability values for multiple degrees of freedom using Greenhouse-Geisser.

## Results

**Behavioral findings.** In order to ensure subjects attended to or ignored the PSS as required, the identification correction rates across three tasks were measured. The identification correction rate was 86% for the VA task, 98% for the NAS task, and 63% for the TAS task, respectively. Statistical results revealed a significantly higher identification correction rate for the NAS task as compared to the TAS task ( $t(21) = -15.059$ ,  $p < 0.001$ ), indicating that detecting the type of PSS is a more demanding task and is assumed to increase attentional load relative to the detection of the number of PSS.

Figure 2 shows the mean magnitude and latency of vocal responses to PSS as a function of task. One-way RM-ANOVAs revealed no



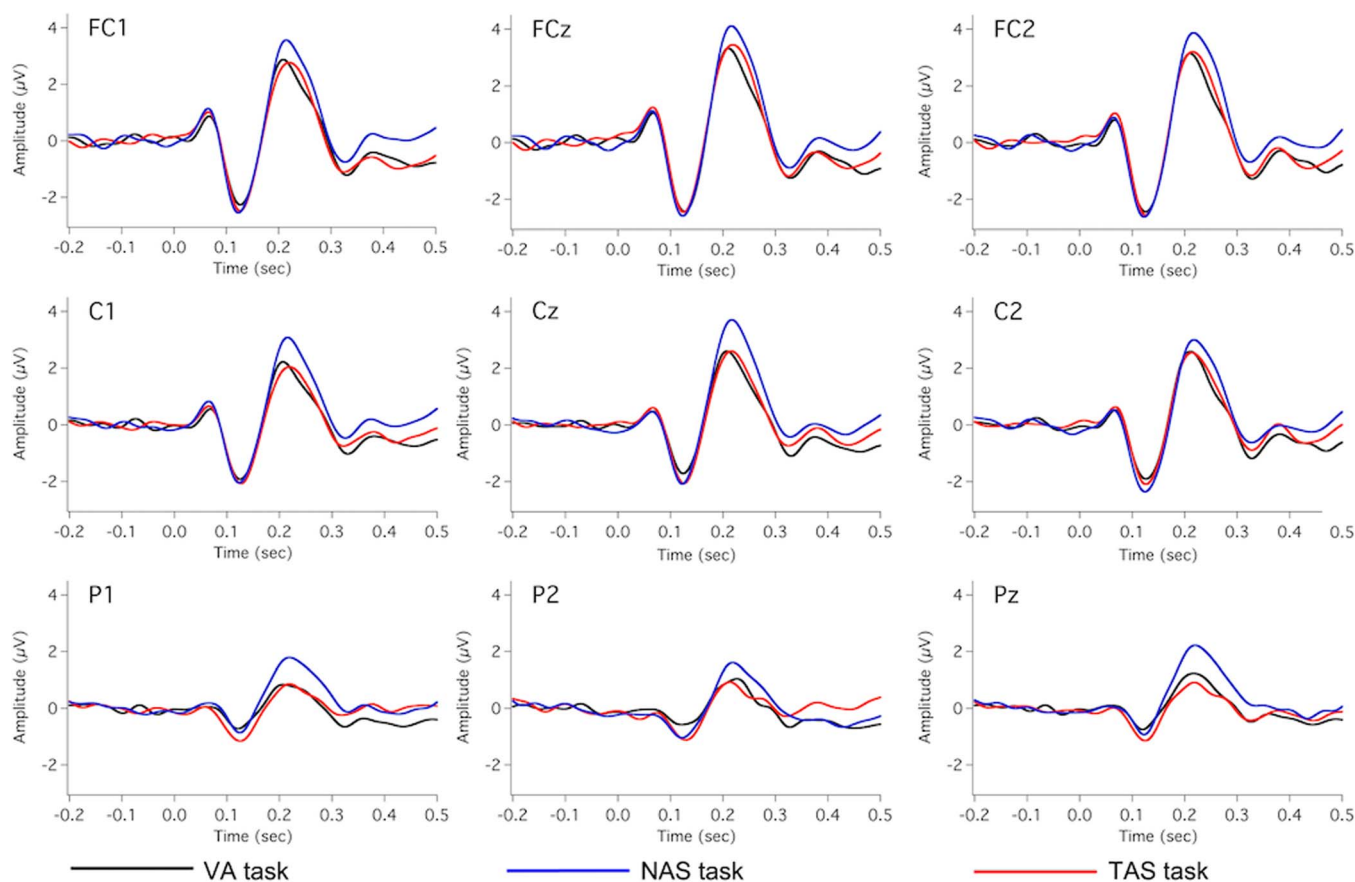
**Figure 2** | T-bar plots (means and standard errors) of magnitudes (A) and latencies (B) of vocal response to pitch perturbations in voice auditory feedback as a function of task.

significant main effect of task on the response magnitude ( $F(2, 40) = 0.213$ ,  $p = 0.809$ ) or latency ( $F(2, 40) = 0.539$ ,  $p = 0.588$ ), indicating that vocal response in magnitude or latency was not modulated as a function of auditory attention.

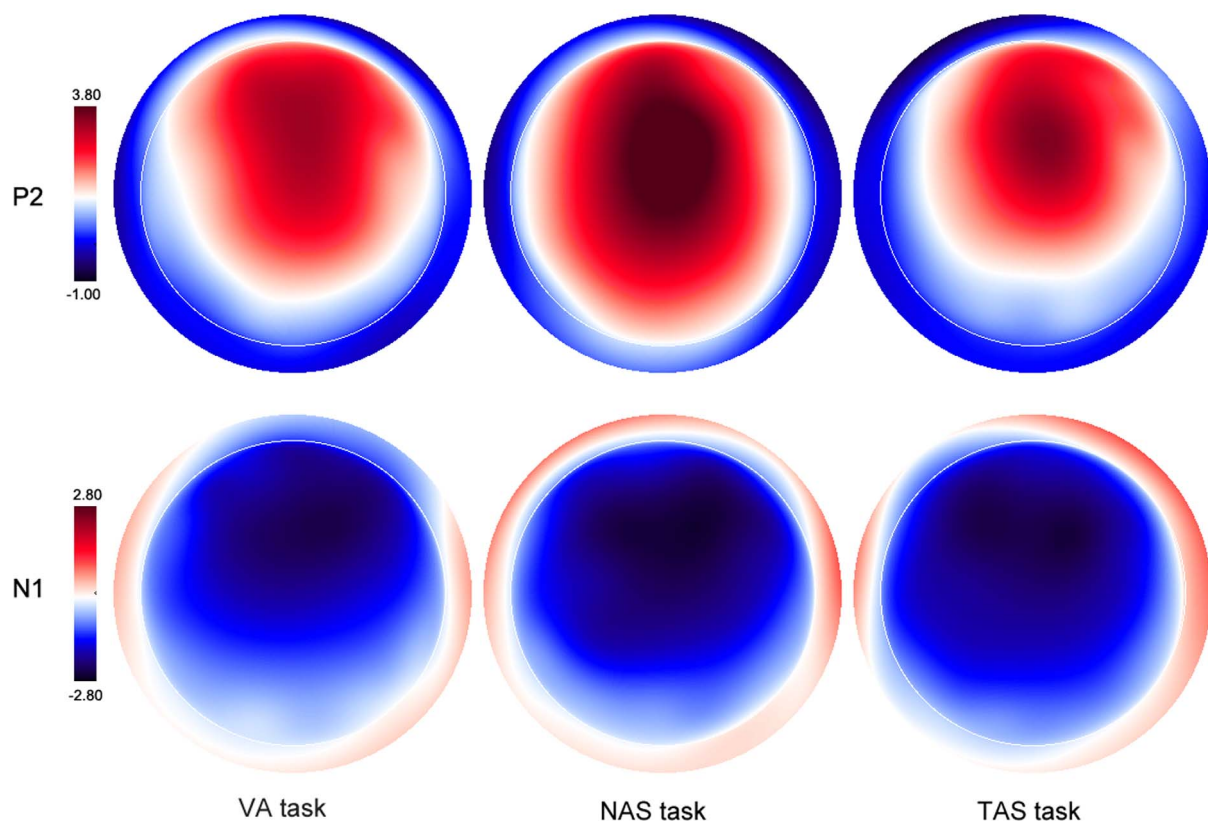
**Neurophysiological findings.** Figure 3 shows the grand-averaged waveforms of ERPs to +200 cents as a function of task. As can be seen, the NAS task elicited greater responses, particularly at the P2 component, than the other two tasks. Similar ERPs were elicited for

the VA and TAS task though. By contrast, N1 amplitude did not differ across three tasks. This task-dependent modulation of neurophysiological response can also be seen in the topographical distributions of the N1-P2 complex (see Figure 4). In addition, the amplitudes of N1 and P2 response appear to be greater in the fronto-central electrodes relative to the parietal electrodes.

One three-way RM-ANOVA of N1 amplitude revealed significant main effects of anteriority ( $F(2, 40) = 25.127$ ,  $p < 0.001$ ) and laterality ( $F(4, 80) = 4.386$ ,  $p = 0.010$ ), whereas main effect of task did not



**Figure 3** | Grand-averaged waveforms of ERPs to +200 cents stimuli in the VA (black lines), NAS (blue lines), and TAS task (red lines), respectively.

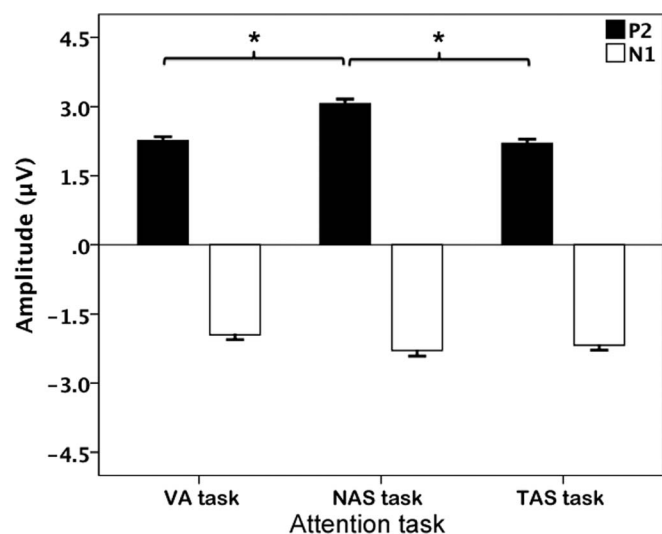


**Figure 4** | Topographical distribution of the grand-averaged ERPs (top: P2 component; bottom: N1 component) to +200 cents stimuli. From left to right are shown the respective ERP distributions for the VA (left), NAS (middle), and TAS task (right).

reach significance ( $F(2, 40) = 1.186, p = 0.316$ ). Post-hoc Bonferroni comparison showed that frontal electrodes elicited the greatest N1 amplitude (absolute value), followed by central and parietal electrodes. There was a significant interaction between anteriority and laterality ( $F(8, 160) = 4.414, p = 0.003$ ), and following-up analyses revealed a significant main effect of laterality only at the frontal sites ( $F(4, 80) = 8.479, p < 0.001$ ). Significant differences of N1 amplitude were found between FC1 and FC3 ( $p = 0.022$ ) and between FC2 and

FC4 ( $p = 0.016$ ), whereas left electrodes did not differ from right electrodes ( $p > 0.05$ ). As for N1 latency, main effects of task ( $F(2, 40) = 0.004, p = 0.996$ ), anteriority ( $F(2, 40) = 4.064, p = 0.057$ ), or laterality ( $F(4, 80) = 0.713, p = 0.501$ ) did not reach significance.

Statistical results of P2 amplitude revealed a significant main effect of task ( $F(2, 40) = 9.392, p < 0.001$ ). Post-hoc Bonferroni comparison revealed that the NAS task elicited significantly greater P2 amplitude than both the VA ( $p = 0.003$ ) and TAS task ( $p = 0.005$ ), whereas no significant difference was found between the VA and TAS task ( $p > 0.05$ ) (see Figure 5). A significant main effect of anteriority ( $F(2, 40) = 59.963, p < 0.001$ ) was also found. Frontal electrodes were associated with greater P2 amplitude than central electrodes ( $p < 0.001$ ) and parietal electrodes ( $p < 0.001$ ), and central electrodes were associated with greater P2 amplitude than parietal electrodes ( $p < 0.001$ ). Main effect of laterality ( $F(4, 80) = 25.576, p < 0.001$ ) also reached significance, which is largely driven by a midline predominance in the response. P2 amplitudes from left and right electrodes, however, did not differ significantly ( $p > 0.05$ ). Regarding P2 latency, there was a significant main effect of task ( $F(2, 40) = 3.869, p = 0.029$ ), in which the VA task produced significantly shorter P2 latency than the TAS task ( $p = 0.042$ ) (211 vs. 221 ms). Main effect of anteriority ( $F(2, 40) = 1.627, p = 0.217$ ) or laterality ( $F(4, 80) = 0.560, p = 0.610$ ), however, did not reach significance.



**Figure 5** | T-bar plots (means and standard errors) of N1 and P2 amplitude as a function of task. The blank and the black bars denote the amplitudes of N1 and P2 components, respectively. The asterisks indicate significant differences between conditions.

## Discussion

The present study aimed to examine whether neurobehavioral processing of pitch feedback errors during the online monitoring of self-produced vocalization can be modulated by auditory attention. Behavioral results revealed no systematic change of vocal response to pitch perturbations irrespective of whether they were ignored or attended. Modulatory effect of auditory attention, however, was observed at the level of cortex. Specifically, P2 amplitude was significantly



increased in response to the attended PSS in the low-load condition when comparing to the unattended PSS in the ignored condition, but had a significant decrease as the attentional load increased through greater task difficulty. Moreover, the attention effect failed to be observed when comparing response to attended PSS in the high-load condition with response to unattended PSS in the ignored condition. These findings demonstrate that auditory-motor integration in voice control can be modulated by auditory attention at the level of cortex. Furthermore, our findings provide supportive evidence for the elastic model of attention that attention effect does not lead to a general enhancement of response to attended targets but occurs as a function of attentional load.

As expected, behavioral results showed that auditory attention had no impact on the modulation of vocal response to PSS. This aspect of our finding is complementary to one recent behavioral study showing that vocal responses observed when vocally-untrained participants were instructed to ignore PSS did not differ from those observed when they were instructed to compensate for PSS that required a focusing attention on the PSS direction<sup>16</sup>, suggesting that vocal compensation for pitch perturbations cannot be consciously modulated. It is also noted from other findings that vocally-untrained participants generally produce compensatory responses to PSS even when told to ignore them<sup>5,14,15</sup>, indicating that they cannot actively suppress their vocal compensation. In line with these findings, the present study provides further evidence that modulation of vocal compensation for pitch perturbations is independent of whether they are attended or ignored, suggesting that auditory-motor integration in voice control appears to be involuntary at the behavioral level and influenced little by attention.

At the level of cortex, however, attention effect was observed as reflected by greater P2 response to the attended PSS in the low-load condition when comparing to the unattended PSS in the ignored condition. This finding provides supportive evidence for our hypothesis that focused attention results in enhanced cortical response in the auditory-motor processing of pitch errors. Despite the methodological difference, this aspect of our finding is in line with neuroimaging and neurophysiological findings of auditory perception showing enhanced brain activity in the auditory cortex<sup>50,51</sup> and greater P2 response<sup>26,29,30</sup> when sounds are attended. More interestingly, P2 response had a significance decrease when the attentional load was increased and returned to the ignored condition, indicating the absence of attention effect in the high-load condition. This finding provides supportive evidence for our hypothesis that cortical responses to attended vs. unattended stimuli may not be generally enhanced but vary as a function of attentional load. This aspect of our finding is in line with the “elastic” effect of auditory attention on the central auditory processing with task load<sup>33,34,52</sup>. Although those studies did not examine change of response to attended sounds as a function of attention load, results from several ECoG studies showed an obvious decrease of positive amplitude of the average ERPs at approximately 170 ms (P170stg) in response to attended stimuli as attentional load increased<sup>33,34</sup>, which is consistent with our findings of decreased P2 response to PSS in the high-load condition.

**Why is cortical response to pitch feedback errors affected by attention?** Regarding the possible mechanisms underlying enhanced P2 responses to the attended vs. unattended PSS in the low-load condition, we postulate that it may result from enhanced response to attended PSS. In light of the gain-based theory of auditory attention<sup>53</sup>, focusing attention improves perceptual detectability to attended features through increasing the gain of neurons that are sensitive to that feature. In the present study, increased gain may be applied to those neurons that are informative for detecting the presence of PSS in voice auditory feedback, leading to enhanced neural activity in response to attended PSS in the low-load

condition. On the other hand, it has been demonstrated that attention to one modality leads not only increased activity of the sensory areas involved in the processing of that modality, but also to suppressed activity in regions associated with other modalities<sup>54–56</sup>. For example, when participants were instructed to attend to melodies and ignore shapes, enhanced activity was observed in the auditory cortex while decreased activity was found in the visual cortex<sup>54,55</sup>. Moreover, there is evidence suggesting that modality-specific selective attention results primarily from decreased processing in the unattended modality rather than increased processing in the attended modality<sup>57</sup>. According to these findings, enhancement effect of attention observed in the present study may result from inhibited processing of unattended PSS in order to free additional resources for the processing of flashing lights in the ignored condition. In addition, we cannot exclude the possibility that changes in cortical responses to attended vs. unattended PSS result from both enhanced response to attended PSS and decreased responses to unattended PSS. Given that different attention conditions were compared with each other rather than a baseline condition (i.e. participants passively listened to the PSS and viewed the lights during vocalization without focused attention) in the present study, however, we cannot determine whether changes in cortical response reflect enhancement in the attended condition, inhibition in the unattended condition, or both. Further studies including a baseline condition should be conducted to verify these speculations.

Interestingly, increased attentional load resulted in the absence of attention effect on the auditory-motor processing of pitch errors, which was primarily driven by a significant decrease of P2 response to attended PSS as load was increased through greater task difficulty. Although speculative, we postulate that decreased cortical response in the high-load condition may be related to the function of working memory. Considerable evidence has demonstrated the role of working memory in top-down attentional control. For example, divided attention requires handling multiple information in working memory relative to selective attention, and the activation in the prefrontal cortex (PFC) that subserves executive functions of working memory<sup>58,59</sup> but decreased activity in the primary sensory areas have been found in the divided attention condition when comparing to the selective attention<sup>54–56</sup>. Moreover, there is evidence showing enhanced activity in the dorsolateral PFC but decreased activity in the precuneus with the attentional load in the visual modality of attention<sup>60</sup>. In light of these findings, it is suggested that there may be an inhibitory influence of working memory on responses in sensory regions by recruiting the PFC. Specifically, the PFC may exert top-down inhibitory control over the primary sensory cortices through function interactions<sup>55</sup> or neural connections<sup>61</sup> between them. Several ERP studies reported supportive evidence to this speculation showing increased auditory-evoked<sup>62</sup> and somatosensory-evoked<sup>63</sup> responses in patients with focal PFC damage, suggesting an inhibitory influence of the PFC on sensory activity in these regions. In the present study, relative to the NAS task, the TAS task demands some sources related to working memory for the processing of pitch perturbation such that the information of the remembered PSS can be maintained and processed (i.e. pitch discrimination to determine whether they were the same or different) for the calculation of the types of all stimuli. Accordingly, it is plausible that decreased cortical responses to PSS in the high-load condition may be caused by inhibited activity of sensory areas due to the involvement of the PFC subserving the function of working memory. Neuroimaging studies needs to be conducted to testify this speculation in the future.

**Differential effects of attention on the N1-P2 complex.** Although previous research on auditory perception has shown enhanced N1 response<sup>25–28</sup> or its magnetic counterpart, the N1m<sup>64,65</sup>, to attended vs. ignored sounds, attention effects on the cortical processing of



pitch errors were only observed in P2 component in the present study. This inconsistency may be related to the difference in the methodology between the present study and others. N1 and P2 responses were recorded in the production of vocal compensation for pitch errors in the present study, which may reflect not only the central auditory processing (e.g. pitch error detection/discrimination) but also the feedback-based motor processing (e.g. pitch error correction). The influence of vocal motor system on central auditory processing has been demonstrated in studies of both animals and humans<sup>17,18,20,22,66</sup>. For example, when comparing with passive listening, utterance onset in the quiet or perturbed condition elicits suppressed brain activity (e.g. N1 or N1m) in the auditory cortex<sup>17,66</sup> whereas perturbations in the middle of utterance causes an increase of ERP response (e.g. P2)<sup>18,22</sup>. It has been suggested that neural generators underlying the N1-P2 complex in correcting for errors in voice auditory feedback may be associated with higher cognitive aspects of vocal output monitoring<sup>18,67</sup>. Thus, differential effects of auditory attention on the N1-P2 complex between the present study and others may be partially attributable to the interaction between auditory and vocal motor system for vocal output monitoring.

On the other hand, this finding provides further evidence of differential mechanisms underlying the generation of N1 and P2 in auditory-motor processing of pitch errors. Generally, neural generators of N1 are mainly located in the primary and secondary auditory cortex<sup>68</sup>. N1 is suggested to reflect the detection of mismatch between the predicted auditory input (i.e. efference copy) and the actual auditory feedback (i.e. corollary discharge) in correcting for feedback errors<sup>18,45</sup>. This speculation is supported by findings that, relative to P2 response, N1 response is more sensitive to the acoustical features of pitch perturbation (e.g. direction or magnitude) but less sensitive to the task demands<sup>45,67,69</sup>. Evidence also suggests that there is a pre-attentive automatic detection of irregularities in sounds in the auditory cortex<sup>24</sup>. If this holds, the earlier stage of cortical processing of unexpected pitch errors may be automatic, which might account for the absence of attention effect on N1 response in the present study. By contrast, P2 receives some contributions from cortical areas in the Sylvian fissure<sup>70</sup>. The posterior Sylvian fissure at the parietal-temporal boundary (area Spt) is proposed to serve as an interface that performs a coordinate transformation between auditory and motor representations<sup>71</sup>. Moreover, there is evidence that P2 amplitude negatively correlated to the magnitude of vocal response<sup>45</sup>, suggesting that change of P2 response may reflect more than simple error detection in voice auditory feedback. Rather, it may reflect the interactions between auditory and motor system that demand higher level of cognitive processes such as attention for auditory processing and vocal motor control, which may be responsible for the modulation of P2 response as a function of attention.

The generality of these observations might be limited by the experimental paradigm. In addition to the PSS, visual stimuli (i.e. the red light) were presented in the ignored condition while they were cancelled in the attended conditions. It is possible that response to attended PSS might be influenced if the visual distraction were presented. It should be noted, however, that our purpose was to examine the effect of attention on response to auditory stimuli no matter whether visual stimuli were presented or not. Despite the methodological difference between the present study and previous studies of selective attention, one common element they share is that attention is directed towards a relevant auditory stimulus and to another stimulus in the ignored condition. Therefore, conclusions from the present findings would not be confounded by this way of attention manipulation.

In conclusion, the present study investigated whether attention can influence the auditory-motor processing of pitch errors during the online monitoring of vocal production. Behavioral results showed no impact of attention on vocal response to PSS irrespective

of whether it is attended. Attention-driven changes of response, however, were observed at the level of cortex. P2 response to attended PSS was significantly enhanced relative to unattended PSS in the condition of low load but returned to the ignored level in the condition of greater load, indicating an enhancement effect of attention in the low-load condition but the absence of this effect in the high-load condition. These findings not only provide further evidence that auditory-motor integration in vocal pitch regulation is involuntary at the behavioral level, but also demonstrate for the first time that attention can modulate the auditory-motor processing of pitch errors as a function of attentional load at the level of cortex.

- Hickok, G., Houde, J. F. & Rong, F. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* **69**, 407–22 (2011).
- Jones, J. A. & Munhall, K. G. Remapping auditory-motor representations in voice production. *Curr. Biol.* **15**, 1768–1772 (2005).
- Larson, C. R. Cross-modality influences in speech motor control: The use of pitch shifting for the study of F0 control. *J. Commun. Disord.* **31**, 489–503 (1998).
- Houde, J. F. & Jordan, M. I. Sensorimotor adaptation in speech production. *Science* **279**, 1213–1216 (1998).
- Burnett, T. A., Freedland, M. B., Larson, C. R. & Hain, T. C. Voice F0 Responses to Manipulations in Pitch Feedback. *J. Acoust. Soc. Am.* **103**, 3153–3161 (1998).
- Chen, S. H., Liu, H., Xu, Y. & Larson, C. R. Voice F0 responses to pitch-shifted voice feedback during English speech. *J. Acoust. Soc. Am.* **121**, 1157–1163 (2007).
- Villacorta, V. M., Perkell, J. S. & Guenther, F. H. Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *J. Acoust. Soc. Am.* **122**, 2306–2319 (2007).
- Bauer, J. J., Mittal, J., Larson, C. R. & Hain, T. C. Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude. *J. Acoust. Soc. Am.* **119**, 2363–2371 (2006).
- Liu, H. & Larson, C. R. Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *J. Acoust. Soc. Am.* **122**, 3671–3677 (2007).
- Liu, H., Wang, E. Q., Verhagen Metman, L. & Larson, C. R. Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson's disease. *PLoS ONE* **7**, e33629 (2012).
- Natke, U., Donath, T. M. & Kalveram, K. T. Control of voice fundamental frequency in speaking versus singing. *J. Acoust. Soc. Am.* **113**, 1587–1593 (2003).
- Chen, Z. *et al.* ERP correlates of language-specific processing of auditory pitch feedback during self-vocalization. *Brain Lang.* **121**, 25–34 (2012).
- Zarate, J. M. & Zatorre, R. J. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage* **40**, 1871–87 (2008).
- Hain, T. C. *et al.* Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Exp. Brain Res.* **130**, 133–141 (2000).
- Patel, S. *et al.* Understanding the mechanisms underlying voluntary responses to pitch-shifted auditory feedback. *J. Acoust. Soc. Am.* **135**, 3036 (2014).
- Keough, D., Hawco, C. & Jones, J. A. Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback. *BMC Neurosci.* **14**, 25 (2013).
- Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M. & Ford, J. M. Fine-tuning of auditory cortex during speech production. *Psychophysiology* **42**, 180–190 (2005).
- Behroozmand, R., Liu, H. & Larson, C. R. Time-dependent neural processing of auditory feedback during voice pitch error detection. *J. Cogn. Neurosci.* **23**, 1205–1217 (2011).
- Niziolek, C. A., Nagarajan, S. S. & Houde, J. F. What Does Motor Efference Copy Represent? Evidence from Speech Production. *J. Neurosci.* **33**, 16110–16116 (2013).
- Eliades, S. J. & Wang, X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* **453**, 1102–1106 (2008).
- Behroozmand, R., Karvelis, L., Liu, H. & Larson, C. R. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clin. Neurophysiol.* **120**, 1303–1312 (2009).
- Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S. & Houde, J. F. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc. Natl. Acad. Sci. USA.* **110**, 2653–8 (2013).
- Nääätänen, R., Paavilainen, P., Rinne, T. & Alho, K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* **118**, 2544–90 (2007).
- Nääätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P. & Winkler, I. “Primitive intelligence” in the auditory cortex. *Trends Neurosci.* **24**, 283–8 (2001).
- Alho, K., Woods, D. L., Algazi, A. & Naatanen, R. Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalogr. Clin. Neurophysiol.* **82**, 356–68 (1992).
- Picton, T. W. & Hillyard, S. A. Human auditory evoked potentials. II. Effects of attention. *Electroencephalogr. Clin. Neurophysiol.* **36**, 191–9 (1974).
- Hillyard, S. A., Hink, R. F., Schwent, V. L. & Picton, T. W. Electrical signs of selective attention in the human brain. *Science* **182**, 177–80 (1973).



28. Maatta, S., Paakkonen, A., Saavalainen, P. & Partanen, J. Selective attention event-related potential effects from auditory novel stimuli in children and adults. *Clin. Neurophysiol.* **116**, 129–41 (2005).
29. Woldorff, M. G. & Hillyard, S. A. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr. Clin. Neurophysiol.* **79**, 170–91 (1991).
30. Neelon, M. F., Williams, J. & Garell, P. C. The effects of auditory attention measured from human electrocorticograms. *Clin. Neurophysiol.* **117**, 504–21 (2006).
31. Pugh, K. R. *et al.* Auditory selective attention: an fMRI investigation. *Neuroimage* **4**, 159–73 (1996).
32. Petkov, C. I. *et al.* Attentional modulation of human auditory cortex. *Nat. Neurosci.* **7**, 658–63 (2004).
33. Neelon, M. F., Williams, J. & Garell, P. C. Elastic Attention: Enhanced, then Sharpened Response to Auditory Input as Attentional Load Increases. *Front. Hum. Neurosci.* **5**, 41 (2011).
34. Neelon, M. F., Williams, J. & Garell, P. C. The effects of attentional load on auditory ERPs recorded from human cortex. *Brain Res.* **1118**, 94–105 (2006).
35. Mottonen, R., van de Ven, G. M. & Watkins, K. E. Attention fine-tunes auditory-motor processing of speech sounds. *J. Neurosci.* **34**, 4064–9 (2014).
36. Zarate, J. M., Wood, S. & Zatorre, R. J. Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia* **48**, 607–618 (2010).
37. Parkinson, A. L. *et al.* Understanding the neural mechanisms involved in sensory control of voice production. *Neuroimage* **61**, 314–322 (2012).
38. Kort, N. S., Nagarajan, S. S. & Houde, J. F. A bilateral cortical network responds to pitch perturbations in speech feedback. *Neuroimage* **86**, 525–35 (2014).
39. Greenlee, J. D. *et al.* Sensory-motor interactions for vocal pitch monitoring in non-primary human auditory cortex. *PLoS ONE* **8**, e60783 (2013).
40. Alho, K., Rinne, T., Herron, T. J. & Woods, D. L. Stimulus-dependent activations and attention-related modulations in the auditory cortex: a meta-analysis of fMRI studies. *Hear Res.* **307**, 29–41 (2014).
41. Degerman, A., Rinne, T., Salmi, J., Salonen, O. & Alho, K. Selective attention to sound location or pitch studied with fMRI. *Brain Res.* **1077**, 123–34 (2006).
42. Rinne, T. Activations of human auditory cortex during visual and auditory selective attention tasks with varying difficulty. *Open Neuroimag. J.* **4**, 187–93 (2010).
43. Hawco, C. S., Jones, J. A., Ferretti, T. R. & Keough, D. ERP correlates of online monitoring of auditory feedback during vocalization. *Psychophysiology* **46**, 1216–1225 (2009).
44. Liu, H., Meshman, M., Behroozmand, R. & Larson, C. R. Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clin. Neurophysiol.* **122**, 951–957 (2011).
45. Scheerer, N. E., Behich, J., Liu, H. & Jones, J. A. ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience* **240**, 176–185 (2013).
46. Boersma, P. Praat, a system for doing phonetics by computer. *Glott International* **5**, 341–345 (2001).
47. Chen, X. *et al.* Sensorimotor control of vocal pitch production in Parkinson's disease. *Brain Res.* **1527**, 99–107 (2013).
48. Li, W. *et al.* Neurophysiological evidence of differential mechanisms involved in producing opposing and following responses to altered auditory feedback. *Clin. Neurophysiol.* **124**, 2161–2171 (2013).
49. Ferree, T. C., Luu, P., Russell, G. S. & Tucker, D. M. Scalp electrode impedance, infection risk, and EEG data quality. *Clin. Neurophysiol.* **112**, 536–44 (2001).
50. Grady, C. L. *et al.* Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport* **8**, 2511–6 (1997).
51. Kawashima, R. *et al.* Selective visual and auditory attention toward utterances-a PET study. *Neuroimage* **10**, 209–15 (1999).
52. Alain, C. & Izenberg, A. Effects of attentional load on auditory scene analysis. *J. Cogn. Neurosci.* **15**, 1063–73 (2003).
53. Hillyard, S. A., Vogel, E. K. & Luck, S. J. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 1257–70 (1998).
54. Johnson, J. A. & Zatorre, R. J. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. *Cereb. Cortex.* **15**, 1609–20 (2005).
55. Johnson, J. A. & Zatorre, R. J. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage* **31**, 1673–81 (2006).
56. Loose, R., Kaufmann, C., Auer, D. P. & Lange, K. W. Human prefrontal and sensory cortical activity during divided attention tasks. *Hum. Brain Mapp.* **18**, 249–59 (2003).
57. Mozolic, J. L. *et al.* Cross-modal deactivations during modality-specific selective attention. *BMC Neurol* **8**, 35 (2008).
58. Curtis, C. E. & D'Esposito, M. The effects of prefrontal lesions on working memory performance and theory. *Cogn. Affect. Behav. Neurosci.* **4**, 528–39 (2004).
59. Petrides, M. The role of the mid-dorsolateral prefrontal cortex in working memory. *Exp. Brain Res.* **133**, 44–54 (2000).
60. Mazoyer, P., Wicker, B. & Fonlupt, P. A neural network elicited by parametric manipulation of the attention load. *Neuroreport* **13**, 2331–4 (2002).
61. Makris, N. *et al.* Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cereb. Cortex.* **15**, 854–69 (2005).
62. Knight, R. T., Scabini, D. & Woods, D. L. Prefrontal cortex gating of auditory transmission in humans. *Brain Res.* **504**, 338–42 (1989).
63. Yamaguchi, S. & Knight, R. T. Gating of somatosensory input by human prefrontal cortex. *Brain Res.* **521**, 281–8 (1990).
64. Fujiwara, N., Nagamine, T., Imai, M., Tanaka, T. & Shibasaki, H. Role of the primary auditory cortex in auditory selective attention studied by whole-head neuroimagnetometer. *Brain Res. Cogn. Brain Res.* **7**, 99–109 (1998).
65. Woldorff, M. G. *et al.* Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc. Natl. Acad. Sci. USA.* **90**, 8722–6 (1993).
66. Houde, J. F., Nagarajan, S. S., Sekihara, K. & Merzenich, M. M. Modulation of the auditory cortex during speech: An MEG study. *J. Cogn. Neurosci.* **14**, 1125–1138 (2002).
67. Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A. & Larson, C. R. Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain Cogn.* **84**, 97–108 (2014).
68. Martin, B. A., Tremblay, K. L. & Korczak, P. Speech evoked potentials: from the laboratory to the clinic. *Eur. Hear.* **29**, 285–313 (2008).
69. Chen, Z., Chen, X., Liu, P., Huang, D. & Liu, H. Effect of temporal predictability on the neural processing of self-triggered auditory stimulation during vocalization. *BMC Neurosci.* **13**, 55 (2012).
70. Crowley, K. E. & Colrain, I. M. A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin. Neurophysiol.* **115**, 732–44 (2004).
71. Hickok, G., Buchsbaum, B., Humphries, C. & Muftuler, T. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* **15**, 673–82 (2003).

## Acknowledgments

This study was funded by grants from National Natural Science Foundation of China (Nos. 31070990, 31371135, 81301675, and 81472154), Guangdong Natural Science Funds for Distinguished Young Scholar (No. S2013050014470), and the Fundamental Research Funds for the Central Universities (No. 13ykzd05). Huijing Hu and Ying Liu contributed equally to this study.

## Author contributions

H.H. and H.L. designed the project; H.H., Y.L. and Z.G. performed the experiment and analyzed the data; H.H., W.L., P.L., S.C. and H.L. interpreted the results; H.H. and H.L. wrote the manuscript; all authors read and approved the final manuscript.

## Additional information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Hu, H. *et al.* Attention Modulates Cortical Processing of Pitch Feedback Errors in Voice Control. *Sci. Rep.* **5**, 7812; DOI:10.1038/srep07812 (2015).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>