Published in final edited form as:

Brain Cogn. 2014 February ; 84(1): 97–108. doi:10.1016/j.bandc.2013.11.007.

Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch

Roozbeh Behroozmand¹, Nadine Ibrahim¹, Oleg Korzyukov¹, Donald A. Robin², and Charles R. Larson¹

¹Speech Physiology Lab, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60208

²Research Imaging Institute, University of Texas Health Science Center San Antonio, San Antonio, TX 78229

Abstract

The ability to process auditory feedback for vocal pitch control is crucial during speaking and singing. Previous studies have suggested that musicians with absolute pitch (AP) develop specialized left-hemisphere mechanisms for pitch processing. The present study adopted an auditory feedback pitch perturbation paradigm combined with ERP recordings to test the hypothesis whether the neural mechanisms of the left-hemisphere enhance vocal pitch error detection and control in AP musicians compared with relative pitch (RP) musicians and non-musicians (NM). Results showed a stronger N1 response to pitch-shifted voice feedback in the right-hemisphere for both AP and RP musicians compared with the NM group. However, the left-hemisphere P2 component activation was greater in AP and RP musicians compared with NMs and also for the AP compared with RP musicians. The NM group was slower in generating compensatory vocal reactions to feedback pitch perturbation compared with musicians, and they failed to re-adjust their vocal pitch after the feedback perturbation was removed. These findings suggest that in the earlier stages of cortical neural processing, the right hemisphere is more active in musicians for detecting pitch changes in voice feedback. In the later stages, the left-hemisphere is more active during the processing of auditory feedback for vocal motor control and seems to involve specialized mechanisms that facilitate pitch processing in the AP compared with RP musicians. These findings indicate that the left hemisphere mechanisms of AP ability are associated with improved auditory feedback pitch processing during vocal pitch control in tasks such as speaking or singing.

Keywords

Absolute Pitch; Relative Pitch; Pitch Processing; Vocal Motor Control; Auditory Feedback; Event-Related Potentials (ERPs)

© 2013 Elsevier Inc. All rights reserved.

Corresponding Author: Charles R Larson, Ph.D., Address: Dept. of Communication Sciences and Disorders, Northwestern University, 2240 Campus Dr Evanston, IL 60208, Telephone: 847-491-2424, Fax: 847-491-4975, clarson@northwestern.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Acquiring musical skills strongly relies on accurate pitch processing for the control of musical instruments or the voice. In recent years, there has been considerable interest in understanding the neural mechanisms that underlie vocal pitch control during speaking and singing (Behroozmand, Karvelis, Liu, & Larson, 2009; Burnett, Freedland, Larson, & Hain, 1998; Chang, Niziolek, Knight, Nagarajan, & Houde, 2013; Eliades & Wang, 2008; Greenlee et al., 2013; F. H. Guenther, Ghosh, S. S. & Tourville, J. A., 2006; Hawco, Jones, Ferretti, & Keough, 2009; Zarate, Wood, & Zatorre, 2010b; Zarate & Zatorre, 2005, 2008). These studies have generally relied on electrophysiological and neuroimaging techniques to unravel information regarding the role of auditory feedback in vocal pitch error detection and correction. The findings of these studies have suggested that vocal pitch control involves interactions between sensory-motor processing networks including but not limited to pre-motor, motor, auditory, parietal and inferior frontal cortices (Chang et al., 2013; Greenlee et al., 2013; F. H. Guenther, 2006; F. H. Guenther, Ghosh, S. S. & Tourville, J. A., 2006; Parkinson et al., 2012; Tourville, Reilly, & Guenther, 2008; Zarate & Zatorre, 2005).

Although it has been established that the functional role of these areas are potentially related to sensory feedback-based control of vocal production, it is not well understood how the specific localization of functions may vary within such complex sensory-motor loops as a result of extensive vocal or musical training or the development of absolute pitch (AP) ability. While many musicians, as a result of years of professional training, can tune instruments relatively accurately to a note, AP musicians have the ability to assign verbal labels to absolute pitches and to identify notes without first determining the note's tonal relationship with another known note (Takeuchi & Hulse, 1993). This unique ability in AP musicians significantly differentiates them from relative pitch (RP) musicians, who control vocal or instrumental pitches using tonal relationships between a target and an external referent note. More interestingly, the duration of musical training is by and large ineffective in differentiating AP and RP musicians, while the early-learning theory suggests that individuals who received musical training before a certain, critical age (3–6 years) may develop AP ability. Data also suggest that AP ability may only develop if said early musical training includes affirming the relation between pitch names and their respective absolute pitches. Takeuchi & Hulse (Takeuchi & Hulse, 1993) argue that if training focuses on the relational aspects of pitch, then the child may not develop AP ability despite receiving musical training within the early-learning period. Adults are therefore unable to learn AP because they are no longer able to perceive a single, absolute tone outside of its relationship to other tones. Regardless, after the critical age, AP ability may not be achieved, meaning that the RP musicians cannot acquire AP ability with increased musical experience (Takeuchi & Hulse, 1993). Such differences in pitch processing between AP and RP musicians suggest that the neural mechanisms of auditory feedback processing for vocal pitch control may differ between these two groups and also between musicians and non-musicians (NM) during speech production and singing.

A study by Schlaug et al. (Schlaug, Jancke, Huang, & Steinmetz, 1995) has reported that AP ability is associated with a larger planum temporale of the left-hemisphere in AP compared with RP musicians. More recently, Jancke et al. (Jancke, Langer, & Hanggi, 2012) provided evidence that AP processing ability may be associated with fewer distant projections and a higher density of peri-sylvian connectivity between different anatomical areas of the brain, including the planum temporale, Heschl's gyrus and lateral superior temporal gyrus. A study by Loui et al. (Loui, Zamm, & Schlaug, 2012) has also found evidence of hyper-connectivity in the left hemisphere areas of the superior temporal gyrus in AP possessors compared with non-AP possessors. Taken together, these reports provide strong support for different neural mechanisms and anatomical connectivity in the left-hemisphere language

areas of the brain in AP musicians. As part of the functional significance of these findings, it should be noted that categorization of musical notes involving mechanisms in the planum temporale is a hallmark of AP ability (Jancke et al., 2012; Loui et al., 2012; Schlaug et al., 1995). However, the question remains if there are other functions of musical or vocal control ability that are associated with AP possessors.

It is well known that the left hemisphere is dominant for speech and language processing and it has been suggested that the AP musicians may utilize these mechanisms in order to gain a higher level of functional expertise for pitch processing during speaking or singing (Jancke et al., 2012; Loui et al., 2012; R. J. Zatorre & Belin, 2001). These functions include fine temporal resolution of acoustical stimuli and the ability to make absolute categorization of acoustical stimuli, such as the ability to perceive the acoustical difference between two vowels. Other than these possibilities, it is still unclear how left hemisphere processing provides some musicians with AP ability. By the same token, it is not known if there are differences in the left and right hemispheres for the control of voice, regardless of its relevance to AP processing in trained musicians.

One potentially useful tool for studying the behavioral and neurophysiological aspects of vocal pitch control is the auditory feedback pitch perturbation paradigm. From the behavioral perspective, it has been shown that delivering pitch-shifts to voice auditory feedback during steady vowel phonations elicit short latency (~100 ms) compensatory vocal responses that stabilize voice fundamental frequency (F0) (Burnett et al., 1998; Chen, Liu, Xu, & Larson, 2007; Donath, Natke, & Kalveram, 2002; Jones & Munhall, 2000; Kawahara & Aikawa, 1996). Such reflexive vocal reactions to feedback pitch perturbation were suggested to be driven by negative feedback controller mechanisms that correct for pitch errors by comparing the voice F0 feedback with an internal representation established by pitch memory or motor-driven mechanisms such as efference copy (Sperry, 1950; von Helmholtz, 1867; Wolpert, Ghahramani, & Jordan, 1995). A disparity between the feedback and predicted output leads to a response that returns F0 to the desired level (Behroozmand et al., 2009; Burnett et al., 1998; Chang et al., 2013; Greenlee et al., 2013; Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Houde, Nagarajan, Sekihara, & Merzenich, 2002; Larson, Altman, Liu, & Hain, 2008).

Evidence for such predictive mechanisms has been found in the audio-vocal modalities by showing that the N1 responses to self-generated sounds or vocalizations are suppressed compared with passive listening to the playback of the same auditory stimuli (Eliades & Wang, 2003, 2005; Heinks-Maldonado et al., 2005; Heinks-Maldonado, Nagarajan, & Houde, 2006; Houde et al., 2002; Müller-Preuss & Ploog, 1981). Although the nature of this phenomenon is not well-understood, it has been proposed that the suppression effect may be driven by motor-induced cancellation of predicted sensory feedback that does not carry new information for the brain (Bendixen, SanMiguel, & Schroger, 2012). This mechanism has also been suggested to be involved in identifying the source of incoming sounds and differentiating self- from externally-generated auditory stimuli (Behroozmand & Larson, 2011; Heinks-Maldonado et al., 2005). However, when the auditory feedback is altered during speaking, the mismatch between predicted and actual sensory input is proposed to enhance neural sensitivity for the detection and correction of vocal motor errors (Behroozmand et al., 2009; Chang et al., 2013; Eliades & Wang, 2008; Greenlee et al., 2013).

In this context, since the auditory feedback signal is hypothetically compared with an internal referent note, it can be expected that trained musicians would utilize different neural mechanisms for voice motor control, and in turn, would show different reactions to pitch perturbations in their voice auditory feedback compared with NMs. It has been demonstrated

that well-trained singers can vocalize and almost completely suppress their vocal responses to perturbations in their voice auditory feedback whereas NMs by contrast produce much larger reflexive responses to such perturbations (Zarate et al., 2010b; Zarate & Zatorre, 2008).

From the neurophysiological perspective, the pitch-shift stimulus serves as a precise temporal marker for a change in the auditory feedback and can be used to study the electrophysiological correlates of vocal pitch error detection and compensation. Several studies have recorded event-related potentials (ERPs) in response to pitch-shifted auditory feedback during vocalization (Behroozmand et al., 2009; Hawco et al., 2009; Heinks-Maldonado et al., 2005). Results of these studies have indicated that different ERP components (e.g. N1 or P2) are sensitive to specific features of pitch-perturbation stimuli while a person is controlling voice F0. For example, the amplitude of the P2 responses was shown to be modulated by the magnitude of pitch-shifts in voice auditory feedback (Behroozmand et al., 2009). In addition, the N1 responses were reported to be sensitive to predictability of the stimulus and whether or not the feedback signal is self- or externally-generated (Bass, Jacobsen, & Schroger, 2008; Behroozmand & Larson, 2011; Heinks-Maldonado et al., 2005; Knolle, Schroger, & Kotz, 2013).

In the present study, we adopted the auditory feedback pitch perturbation paradigm combined with ERP recordings in order to investigate the behavioral and neurophysiological aspects of vocal pitch control in three subject groups of AP and RP musicians as well as NMs. Subjects were asked to repeatedly maintain a steady vocalization of the vowel sound / a/ while they randomly received brief (200 ms) upward (+100 cents) and downward (−100 cents) pitch-shifts in their voice auditory feedback. The recording of the ERPs in response to feedback pitch perturbations allowed us to specify the spatio-temporal characteristics of the ERP responses during vocal pitch control, specifically in terms of modulation of individual ERP components by stimulus features across different subject groups. Furthermore, the electrophysiological correlates of lateralized functions during vocal production and control were studied by comparing the ERP responses in the left and right hemispheres. We hypothesized that the AP musicians would show greater left-hemisphere activation, as evidenced by larger magnitudes of the P2 component, compared with RP and NM subjects. In addition, we predicted that neural mechanisms of training-induced enhancement of audio-vocal integration in musicians would enable them to control their vocal pitch more accurately compared with the NM group.

2. Materials and Methods

2.1. Participants

This study used a sample of 34 (15 male and 19 female) subjects from Northwestern University, ages 18–25. Each subject had normal hearing and no history of neurological or speech disorders. In this population sample, 23 musicians (12 with RP and 11 with AP ability) were recruited from the Bienen School of Music, while the remaining 11 untrained NM subjects (control group) were recruited from general Northwestern student population. The musician subjects were categorized based on their self-identification as RP or AP possessors and we further confirmed this identification with the battery of musical tests described below. The musicians played instruments including guitar, piano, violin, cello, clarinet, saxophone, trombone, trumpet, tuba, bassoon, French horn, oboe and flute. Eight out of twelve RP musicians and eight out of eleven AP musicians were formally trained vocalists and all of the remaining musicians in both groups had received vocal training at some point during their aural skills education. Musicians in RP and AP groups were similar in terms of the onset age and nearly identical in terms of duration of musical training. The average onset age of musical training for RP and AP musicians was 9.09 (range: 4–16) and

6.73 years (4–11), respectively. The average duration of musical training for RP and AP musicians was 11.64 (range: 4–17) and 12.23 (range: 7–16) years, respectively. All subjects in both groups started their musical training at the age of 11 or younger except for two subjects in the RP group that started music at the age of 14 and 16. The average onset age of RP group was 7.77 years without these two subjects. Although these two subjects started their training later, results of the musical test scores, vocal behavior and ERPs did not show any difference between their performances compared with the RP group, and therefore, they were categorized as RP musicians.

2.2. Musical Proficiency Test

With the help of Aural Skills and Music Professionals in the Bienen School of Music at Northwestern University, a battery of four tests was created to corroborate a musician's self-identification as possessing either RP or AP ability. The tests were administered in order to classify their ability in terms of pitch perception, identification, discrimination and production. NMs took this test despite a lack of musical training to function as a control. Their data highlighted the difference between trained RP and AP musicians as well as the significance of vocal and musical training. Administering this test to NMs also ensured that none were especially musically gifted in order to avoid potential confounds related to their musical ability.

Test 1 (Individual Chromatic Pitches)—This test was designed to test subjects' ability to accurately reproduce and identify pitches delivered via piano notes. AP musicians, in addition to reproducing pitches, should also be able to accurately recognize a pitch and assign it a verbal label (e.g., B $\flat$, A, C $\sharp$). The test was therefore split into two scored parts that assess basic vocal (Part A) and pitch-identification ability (Part B). Since even NMs can perform a basic reproduction task, Part A was meant to rule out those subjects who were unable to do so. Due to the possibility of RP or AP musicians lacking the ability to vocally produce notes, Part A functioned as an exclusionary filter; we aimed to test musicians who, regardless of instrument of training, possessed strong vocal skills reinforced by higher musical education. Part B tested subject's ability to hear a pitch and subsequently assign to it a verbal label without relative clues. The test's musical range guaranteed subjects would encounter pitches that fell both within and outside their singing range. Additionally, the variety of octave spanning note intervals eliminated the sense of pattern that RP musicians rely on to identify pitches. RP musicians may have correctly labeled the note but would not have done so with the same accuracy and precision as AP musicians. Accurate pitch identification is a hallmark of AP ability; therefore, the scoring on this test was critical to the ultimate confirmation of a subject's identification as possessing AP or RP ability.

Test 2 (Tonal Sight-Singing)—In this test, subjects were asked to sight sing along a sequence of notes after hearing the starting note, without first studying the piece. This test emphasized step-wise motion and interval maintenance and the goal was to produce the last not correctly relative to the starting note in the beginning of the sequence. The frequency of the final note was scored for accuracy. The piece in this test embodied a strong sense of tonality and built-in elements made it easy to accurately sing and therefore end on the correct frequency. This piece both began and ended on A 440/220 Hz. AP musicians, regardless of common errors in interval-maintenance, were expected to end on the correct frequency because of the AP ability to produce notes without relative clues. RP musicians, however, were expected to either sing it accurately, using elements in the piece to correct interval errors, or would make an interval mistake that would then be reflected in the ending note frequency. While all AP musicians were expected to attain a perfect score on this test, only some RP musicians were expected to do so.

Test 3 (Atonal Sight-Singing)—This test was also a sight-singing task but in contrast with Test 2 which was tonal and chromatic, the piece in this test had no tonal center. This test utilized a wide variety of intervals and included few short snippets of areas with recognizable scalar structures. While this test contained some repetitive harmonic elements, they were subtle. Far more than in Test 2, if a subject incorrectly sang an interval in this test, they would be less able to adjust using context clues and their final note would reflect the error. Test 3 illuminated the difference between those who relied completely on their aural skills and ear training but finished on the wrong note and those who used an internal pitch representation to self-correct at the end, regardless of mistakes made throughout the piece. The former describes a typical RP musician and the latter an AP musician. It is important to note that while an RP musician may have correctly sight-sung the test and ended up on the right note despite the difficulty level, AP musicians would have done so much more commonly. For musicians who ended on the correct note, their score on Test 1 Part B was to be taken into consideration to identify AP.

Test 4 (Individual Microtonal Pitches)—In this test, subjects were asked to identify microtonal pitch variations in a series of individual notes. They listened to the playback of each note that was either in tune or deviated (maximum ± 50 cents) from the original note given on the sheet music. They were asked to listen to each note and determine a) if it is in-tune or not, and b) if it is not in tune, whether it is flat or sharp. Subjects were also asked to determine the degree to which the note was out of tune (“very flat”, “a little flat”) but this part was not scored due to irreconcilable differences in how musicians perceive “very” versus “a little” flat. This test was meant to exacerbate the results of Test 1 Part B; results from Test 4 allowed analysis of any potential range issues that may have presented themselves as an inability to identify a pitch in Test 1 Part B (the vocal range was similar in the two tests). In addition, AP musicians were expected to perform better than RP musicians despite the unique difficulty that atonal music presents AP possessors. Because AP is not a uniform ability among all possessors, this test also made it possible to distinguish between those with very accurate AP and those who, despite having AP, would become less accurate once outside of their range due to self-doubt. The score of this test cannot be judged alone as an indicator of RP or AP, but rather serves to deepen the understanding of a musician’s other scores.

The subjects’ performance was evaluated and they received a score between 0 to 100% based on their pitch perception, identification, discrimination and production accuracy in each test. The mean and standard deviation of NM, RP and AP scores are reported in table 1 for each individual test and a composite score is calculated for each group as the mean of the scores in all four tests. Objective analysis of these scores showed that the musicians’ performance (both AP and RP) was significantly better than the NM group in all four tests (two-sample t-test, $p < 0.001$). In addition, musicians in the AP group scored significantly higher (two-sample t-test, $p < 0.05$) than the RP group on Part B of Test 1 (individual chromatic pitch) and Test 3 (atonal sight singing). These scores confirmed our predictions related to improved musical performance in trained musicians and the results of the musical proficiency Tests 1 (Part B) and 3 helped us further verify the self-reported AP or RP ability of the recruited musicians.

2.3. Experimental Design

The experiment consisted of two blocks of trials during which subjects were asked to sustain a steady vocalization of the vowel /a/ for approximately 2 seconds at their conversational pitch and loudness. This vocal task was repeated 120 times during each block while subjects took short breaks (1–2 seconds) between successive utterances. During each vocalization trial, subjects were presented with one brief (200 ms duration) randomly chosen upward

(+100 cents) or downward (−100 cents) pitch shift stimulus in their voice auditory feedback. The onset time of these mid-utterance pitch-shift stimuli was randomized between 500–1000 ms after the vocalization onset in each trial. The rising and falling duration of the change in pitch was approximately 16 ms, which is perceptually recognized as instantaneous. The total duration of each block of trials was approximately 10–15 minutes and the subjects were given 2 minutes break between the two blocks.

2.4. Voice Acquisition System

Subjects were seated in a sound-treated room in which their voice was picked up with an AKG boomset microphone (model C420), located about 1" from the mouth and amplified with a Mackie mixer (model 1202-VLZ3). The onset of the vocalization was detected using a voice onset detector (VOD) module implemented in Max/Msp (Cycling 74, v.5.0). The Max/Msp program also controlled an Eventide Eclipse Harmonizer to randomly present an upward (+100 cents) or downward (−100 cents) pitch shift stimulus in the middle of each vocalization, with the onset time randomized between 500–1000 ms after the vocalization onset. A TTL pulse was also generated by the Max/Msp to mark the onset of mid-utterance pitch shift stimuli in each trial, and this TTL pulse was used for time-locked averaging of neural and behavioral responses to pitch perturbations in voice auditory feedback.

Voice, feedback and TTL pulses were sampled at 10 kHz using PowerLab A/D Converter (Model ML880, AD Instruments) and recorded on a laboratory computer utilizing Chart software (AD Instruments). Measurements from the recorded signals in a number of trials indicated an average time delay of approximately 8 ms between the voice and its auditory feedback. Subjects maintained their conversational F0 levels with a voice loudness of about 70–75 dB, and the pitch-shifted feedback was delivered through Etymotic earphones (model ER1-14A) at about 80–85 dB. The 10 dB gain between voice and feedback channels was used to partially mask air-born and bone-conducted voice feedback. Pilot testing confirmed that the 10 dB gain in the voice feedback was sufficient to mask the air-born voice signal, and the sound-attenuated booth eliminated external noise interference. A Brüel & Kjær sound level meter (model 2250) along with a Brüel & Kjær prepolarized free-field microphone (model 4189) and a Zwislocki coupler were used to calibrate the gain between voice and feedback channels.

2.5. ERP Acquisition and Analysis

The electroencephalogram (EEG) signals were recorded from 32 sites on the subject's scalp using an Ag-AgCl electrode cap (EasyCap GmbH, Germany) following the extended international 10–20 system (Oostenveld & Praamstra, 2001) including left and right mastoids. Recordings were made using the average reference montage in which outputs of all of the amplifiers are summed and averaged, and this averaged signal is used as the common reference for each channel.

Scalp-recorded brain potentials were low-pass filtered with a 400 Hz cut-off frequency (anti-aliasing filter) and then digitized at 2 KHz and recorded using a BrainVision QuickAmp amplifier (Brain Products GmbH, Germany) on a computer utilizing BrainVision Recorder software (Brain Products GmbH, Germany). Electrode impedances were kept below 5 k Ω for all channels. The electro-oculogram (EOG) signals were recorded using two pairs of bipolar electrodes placed above and below the right eye and on the lateral canthus of each eye to monitor vertical and horizontal eye movements. An electrode with unipolar shielded cable was placed on the subjects' forehead and was connected to the QuickAmp dedicated slot for ground connection.

The BrainVision Analyzer 2 software was used to analyze recorded EEG signals in order to calculate ERPs time-locked to the onset of mid-utterance upward and downward pitch-shift stimuli that were marked by the TTL pulses. The recorded EEG was filtered offline using a band-pass filter with cut-off frequencies set to 1 and 30 Hz (−24dB/oct) and then segmented into epochs ranging from −100 ms before and 500 ms after the onset of the stimulus. Following segmentation, artifact rejection was carried out by excluding epochs with EEG or EOG amplitudes exceeding $\pm 50 \mu\text{V}$. Individual epochs were then subjected to baseline correction by removing the mean amplitude of the 100 ms-long pre-stimulus time window for each individual EEG channel and separately averaged for upward and downward stimuli across all trials for each individual subject. A minimum number of 100 trials was used to calculate the grand averaged ERP responses in all subjects.

2.6. Voice Response Analysis

The pitch frequency of the recorded voice signals was extracted in Praat (Boersma, 2001) using an autocorrelation method and then exported to MATLAB (Mathworks, Inc.) for further processing. The extracted pitch frequencies were segmented into epochs ranging from −100 ms before to 1000 ms after the onset of pitch-shift stimuli marked by the TTL pulses. The pitch frequencies were then converted from Hertz to Cents scale using the formula: $\text{Cents} = 1200 \times \log_2(F_2/F_1)$ in which F_1 is the baseline (pre-stimulus) pitch frequency from −100 to 0 ms and F_2 is the post-stimulus pitch frequency from 0 to 1000 ms. The calculated pitch contours in Cents were averaged across trials separately for each stimulus direction in individual subjects. Following analysis for individual subjects, grand averages of the vocal responses were calculated across the NM, RP and AP subject groups. The magnitude and latency of vocal responses to pitch-shift stimuli were calculated by finding the most prominent peak in 200 ms -long time windows centered at 250 ms. The range of these time windows were selected based on finding the peak of the grand averaged vocal responses and also visual inspection of the waterfall plots of the single trial contours across all subjects in each group.

3. Results

3.1. Vocal Responses to Feedback Pitch Perturbation

The primary finding from the behavioral data is that, compared with the NM group, the AP and RP musicians produced faster vocal responses to pitch-shifts in their auditory feedback and they were able to readjust their voice F_0 and return to the baseline after the feedback pitch perturbation was removed. All three groups of subjects showed compensatory vocal responses to feedback pitch perturbation (figure 1). It is apparent that the peak of the vocal response in NMs occurred at a longer latency and its magnitude was smaller for downward pitch shifts compared with the RP and AP musicians. Boxplot representations of vocal response peak latencies and magnitudes are provided in figure 1 for each stimulus direction separately. It is also evident that the NMs failed to return to their voice baseline pitch only in response to upward pitch-shifts, whereas for downward shifts, all subjects returned their voice pitch to its baseline frequency.

The magnitude and latencies of the vocal response peaks were analyzed using a Mixed-ANOVA with the group as a between-subjects factor and the stimulus direction as a within-subjects factor (repeated measure). Results indicated no main effects for peak magnitudes but a significant main effect of subject group was found for the peak latencies ($F(2,31)=6.09$, $p=0.006$). Post-hoc tests using Bonferroni's correction showed that this main effect was due to a significantly shorter latency of vocal responses to downward pitch-shift stimuli in AP ($p=0.011$) and RP ($p=0.007$) musicians compared with NMs (see the boxplots in figure 1b). Furthermore, ANOVA tests were performed for each stimulus direction

separately to test for the group effect in non-overlapping 5 ms bins throughout the whole post-stimulus time window. These tests were corrected for multiple comparisons using Bonferroni's method accounting for the total number of groups, stimulus directions and time bins. Results indicated that the magnitude of vocal responses were significantly different ($p < 0.001$) in NM compared with AP and RP groups in response to upward stimuli at latencies above 485 ms. The time points of significant difference are marked with "*" in figure 1a. Moreover, the magnitude of the vocal responses in the last bin within the post-stimulus time window (latency: 1000 ms) was compared with that of the peaks for both stimulus directions across AP, RP and NM groups. Results revealed a significant difference ($p < 0.001$) for both AP and RP musicians in response to upward and downward stimuli, but for the NMs, these responses were significantly different ($p < 0.001$) only in response to downward but not upward pitch-shift stimuli. This finding demonstrates that in contrast with the AP and RP musicians, the NM group failed to re-adjust their voice F0 and did not return to the baseline pitch after the offset of the upward pitch-shift stimulus (see figure 1a).

3.2. Electrophysiological Responses to Feedback Pitch Perturbation

The primary finding of the recorded ERPs is that the AP musicians produced larger amplitude P2 responses to pitch-shifted voice feedback in the left hemisphere compared with the RP and NM groups. The grand averaged ERPs for the NM, RP and AP groups are illustrated for one representative contact on the left hemisphere (T7), right hemisphere (T8) and vertex (Cz) in response to upward and downward pitch-shifts in figures 2a and 3a, respectively. These contacts were chosen based on visual inspection of the topographical distribution maps plotted in figures 2b and 3b and also identifying the sites with largest magnitude responses across all subjects. The N1 components were characterized by a negative deflection at about 150 ms post-stimulus over the vertex (Cz electrode) that inverts in polarity (positive deflection) over the temporal lobe areas (e.g. electrodes T7 and T8). In contrast, the P2 appeared as a positive deflection at about 250 ms over the vertex electrode that inverts (negative) over the bilateral temporal sites. As can be seen in figure 2a, the N1 responses to upward pitch-shifts do not appear to be different across NM, RP and AP groups at vertex, whereas for the P2 potentials, the AP musicians seem to generate larger amplitude responses than those generated by the NM and RP groups. However, for downward pitch-shifts, the N1 appears to be larger at the vertex for AP and RP musicians compared to the NM group and a similar trend in the data is observed for the P2 potentials with stronger responses in the AP musicians compared with the RP and NM groups (figure 3a).

The significance of the differences in the ERP responses to upward and downward pitch shifts across AP, RP and NM groups were statistically assessed by extracting the mean amplitude of the grand averaged N1 and P2 responses in a 20 ms-long time window centered at 150 and 250 ms post-stimulus (see the highlighted areas at vertex in figures 2a and 3a). The time windows were chosen based upon the point of the greatest response magnitude over the vertex (Cz) electrode. The peak latencies of the N1 and P2 responses were also extracted by finding the largest negativity (N1) and positivity (P2) within a 100 ms-long time window centered at 150 and 250 ms. For the vertex site, the mean amplitude and latency of the N1 and P2 responses were analyzed using a Mixed-ANOVA with the group (NM, RP and AP) as a between-subjects factor and stimulus direction (up vs. down) and frontality (Fz: frontal, FCz: fronto-central, Cz: Central and CPz: centro-parietal) as within-subjects factors (repeated measures). For the left and right hemispheres, responses were analyzed for four contacts on each side using a Mixed-ANOVA with the group as a between-subjects factor and stimulus direction, laterality (left: M1, FT7, T7, TP7 vs. right: M2, FT8, T8, TP8 electrodes) and frontality (M: Mastoid, FT: fronto-temporal, T: temporal and TP: temporo-parietal) as within-subjects factors. The four contacts over the vertex left hemisphere and right hemisphere were chosen based upon visual inspection of the

topographical distribution maps of the N1 and P2 components (figures 2b and 3b) and identifying the areas with the largest response amplitudes.

Results of the analysis for the N1 responses at vertex indicated significant main effects of group ($F(2,31)=6.47$, $p=0.004$), direction ($F(1,62)=7.39$, $p=0.007$) and frontality ($F(2,62)=14.59$, $p<0.001$). Post-hoc tests using Bonferroni's correction for multiple comparisons showed that the vertex N1 responses were significantly larger in the AP ($p=0.003$) and RP ($p=0.011$) musicians compared with the NM group only for the downward stimuli (figure 3a). In addition, the stimulus direction significantly modulated the N1 responses only in the AP ($p=0.021$) and RP ($p=0.018$) musicians with larger responses to downward compared with upward pitch-shifts. The N1 responses were significantly larger ($p=0.007$) at FCz and Cz compared with Fz and CPz electrodes and therefore had a fronto-central distribution. The analysis of the peak latency of the N1 responses at vertex did not show any significant main effect.

The analysis of the N1 potentials on the left and right hemispheres over the bilateral temporal lobe areas revealed significant main effects of laterality ($F(1,93)=14.06$, $p=0.001$) and frontality ($F(3,93)=19.32$, $p<0.001$) along with a significant group $\times$ laterality interaction ($F(2,93)=6.62$, $p=0.004$). The N1 responses had a temporo-parietal distribution with significantly larger ($p=0.001$) responses at temporo-parietal and mastoid compared with temporal and fronto-temporal contacts. The group $\times$ laterality interaction was examined using post-hoc tests and the results indicated that the N1 responses were left-lateralized for both stimulus directions only in the NM group ($p=0.015$) but not in the AP and RP musicians (see the N1 barplots in figures 2c and 3c). Furthermore, analysis of the N1 responses separately on the left and right hemispheres revealed that the right-hemisphere N1 responses were significantly larger in the AP ($p=0.003$) and RP ($p=0.001$) musicians compared with the NM group. No significant main effect was revealed for the peak latency of the N1 responses over the bilateral temporal lobe areas.

For the P2 component at vertex, results showed only a significant main effect of group ($F(2,31)=5.71$, $p=0.008$) and frontality ($F(2,62)=12.09$, $p<0.001$). Post-hoc tests indicated that the P2 responses were significantly larger in the AP musicians compared with RP ($p<0.001$) and NM ($p<0.001$) groups for both upward and downward pitch-shift stimuli but no difference was found between P2 responses in the RP and NM groups. The P2 responses had a central distribution and were significantly larger ($p=0.014$) at Cz compared with Fz, FCz and CPz electrodes. The analysis of the peak latency of the P2 responses at vertex did not show any significant main effect.

Analysis of the P2 responses on the left and right hemispheres over the bilateral temporal lobe areas revealed significant main effects of subject group ($F(2,31)=5.14$, $p=0.012$) and laterality ($F(1,93)=5.38$, $p=0.027$) along with a significant group $\times$ direction interaction ($F(2,93)=4.78$, $p=0.013$). Post-hoc tests indicated that the AP musicians had bilaterally stronger P2 responses compared with the NM ($p=0.004$) and RP ($p=0.009$) groups for upward pitch-shifts (figure 2c). However, for downward stimuli the P2 potentials were significantly stronger for the AP compared with RP ($p=0.025$) and NM ($p=0.006$) groups only in the left hemisphere (figure 3c). In addition, the left-hemisphere P2 responses were also significantly stronger ($p=0.017$) for the RP musicians compared with the NM group. Furthermore, for upward pitch-shifts, the P2 potentials were right-lateralized only in the NM group ($p=0.018$) and for downward stimuli the right-lateralization was significant for both the NM ($p=0.011$) and RP ($p=0.015$) groups (see the P2 barplots in figures 2c and 3c). The P2 responses had a temporal distribution with significantly larger ($p=0.029$) responses at temporal and temporo-parietal compared with fronto-temporal and mastoid contacts. No

significant main effect was revealed for the peak latency of the P2 responses over the bilateral temporal lobe areas.

3.3. Neural and behavioral response correlation

We examined the relationship between neural and behavioral responses to pitch perturbations in voice auditory feedback by calculating the correlation coefficients between the ERP amplitudes and the magnitude and latency of the compensatory vocal responses using Pearson's method. Results of the analysis revealed a negative correlation between the amplitude of the P2 responses and the latency of behavioral vocal compensations only in response to downward (−100 cents) pitch-shift stimuli. This negative correlation was significant for P2 ERP responses at Fz ($r = -0.418$, $p < 0.05$), FCz ($r = -0.512$, $p < 0.01$) and Cz ($r = -0.546$, $p < 0.01$) electrodes, indicating that larger P2 amplitudes were correlated with shorter latency (faster) vocal responses to feedback pitch perturbations. No significant correlation was found between ERP amplitudes and vocal response latencies for upward pitch-shift stimuli. In addition, there was no significant correlation between the ERP amplitudes and vocal response magnitudes for either upward or downward pitch shifts. The correlation plots for the P2 amplitudes at Cz electrode and the latencies of the vocal response to upward and downward pitch shifts are shown in figures 4a and 4b, respectively.

4. Discussion

In the present study, we investigated the modulatory effect of musical training on the audio-vocal cortical networks that mediate voice pitch control using auditory feedback. Our data provide behavioral and electrophysiological evidence that trained musicians demonstrate improved abilities for tracking pitch variations in their voice auditory feedback compared with NMs. In addition, our findings showed that the neural processing of pitch perturbation in voice auditory feedback is further enhanced in musicians who possess AP ability compared to musicians with RP or NMs. Even though the neural mechanisms underlying AP ability remain elusive, findings of the present study suggest that this enhanced pitch processing capability may have resulted from the development of specific left-hemisphere neural mechanisms that are specialized for providing a strong internal neural representation of pitch in AP musicians. This conclusion is based upon our data showing that the P2 responses to pitch-shifted voice feedback in the left hemisphere were significantly stronger in AP compared with RP musicians. This effect is more evident for P2 responses to downward pitch-shift stimuli where the AP musicians had significantly larger P2 responses compared with RP and NM groups in the left-hemisphere (see figure 3). The left hemisphere P2 responses were also larger in RP musicians compared with the NM group, suggesting that this component may reflect some levels of pitch processing enhancement in musicians compared with NMs, and in addition can be used as an electrophysiological marker of AP ability in trained musicians.

The models of audio-vocal integration (Hickok, Houde, & Rong, 2011; Rauschecker, 2011; Rauschecker & Scott, 2009) have proposed a dual neural processing mechanism that consists of posterior dorsal (PD) and antero-ventral (AV) pathways. In this model, the AV pathway consists of projections that transform auditory representations to motor commands for driving pre-motor and motor cortex during speech production (Hackett, Stepniewska, & Kaas, 1998), and the PD pathway provides an interface between pre-motor and auditory cortices that allows for comparison of auditory feedback information with internally predicted motor commands. While the AV and PD pathways may serve as a basic interface for feedback-based vocal monitoring and control, it has been shown that the anterior cingulate cortex (ACC), bilateral STG, insula and left putamen are more strongly recruited by the trained musicians during voice pitch control, suggesting that these areas may develop to form a high-level neural circuitry for feedback monitoring as a result of musical practice

(Zarate et al., 2010b; Zarate & Zatorre, 2008). Patterson et al. (Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002) have proposed a system for the processing of musical stimuli that projects from the brainstem up through Heschl's gyrus, lateral superior temporal sulcus and planum temporale, with growing complexity of processing and extraction of melody in higher cortical areas bilaterally. At higher cortical areas, there may also be greater asymmetric processing involving the right hemisphere (R. Zatorre, Belin, & Penhune, 2002). The enhancement of the right-hemispheric N1 potentials in RP and AP musicians compared with the NM group in the present study (figures 2 and 3) may reflect neural activity in higher-level networks within the non-primary auditory cortex areas that are involved in extraction of the significance of the pitch-shifted feedback for vocal control. It is possible that the musicians need to recruit these neural circuits in the right hemisphere to access memory representation stored in the non-primary auditory cortex for stimulus matching and pitch labeling. In addition, for the P2 potentials, the bilateral enhancement in response to upward and left-hemisphere enhancement in response to downward pitch shifts in musicians in the present study may be an indication that the neural processing is further shifted to an even higher-level audio-vocal network in the ACC, insula and putamen that develop as a result of extensive musical practice (Zarate, Wood, & Zatorre, 2010a).

The identification of lateralized ERP responses in the present study can be discussed in the context of previous reports on anatomical and functional neural asymmetries for sound processing in AP musicians. Schlaug et al. (Schlaug et al., 1995), using volumetric measurement of anatomical structures, demonstrated a leftward asymmetry in the size of the planum temporale in AP musicians compared with RP musicians and NMs. Functional imaging (fMRI) and PET measurements during musical sound presentation have also shown that AP musicians exhibit stronger activation of the left superior temporal gyrus (STG) and posterior dorso-lateral frontal cortex (PDLFC) compared with RP musicians (Loui et al., 2012; R. J. Zatorre, Perry, Beckett, Westbury, & Evans, 1998). These results suggest that the functional capacity of the cortex subserving enhanced pitch processing associated with AP ability may increase with development of a stronger left-hemisphere activation in specific brain areas possibly related to musical training during a critical period of early brain development. This effect is reflected through the increased response amplitude of the P2 component in the present study and can be considered as evidence for specialization of left-hemisphere mechanisms for the development of AP ability in trained musicians. While it is clear that left hemisphere mechanisms of musical ability emerge in musicians with AP ability, it is still likely that many of the right hemisphere mechanisms and locations that have previously been described retain their importance as well (Zarate et al., 2010b; Zarate & Zatorre, 2008).

Across group comparison of the ERP components in different conditions revealed several effects that provide important information for understanding pitch processing mechanisms in trained musicians and NMs during vocal production and control. First, our data propose that musical training enhances the neural sensitivity for downward (decreasing) pitch changes in voice auditory feedback. Results indicated that both the AP and RP musicians showed larger N1 responses to downward compared with upward pitch-shifts (stimulus direction effect), suggesting greater neural sensitivity to decreasing feedback pitch frequency during vocalization (compare vertex N1 responses in figures 2b and 3b). However, the P2 components were not modulated by stimulus direction, suggesting that the detection of pitch change direction occurs during earlier stages of neural processing, as shown by the differences in the N1 potential. It has recently been suggested that the posterior superior temporal sulcus interacts with the intraparietal sulcus to encode direction of pitch shift stimuli (Zarate et al., 2010b). This network could be more highly developed in the musicians with AP and RP ability and therefore could be a contributor to the enhanced N1 responses for downward pitch-shift stimuli in the present study. The reasons why the ERP responses

were stronger for downward compared with upward pitch shifts are not clear, but it is possible that the tendency for musicians to go 'flat' or to reduce the F0 of their instrument may be more salient than going 'sharp'.

The results of the behavioral responses to stimulus direction can be discussed with regard to peripheral mechanisms of voice control. In a study by Liu et al. (Liu, Behroozmand, Bove, & Larson, 2011) it has been reported that the cricothyroid (CT) and thyroarytenoid (TA) muscles of the larynx show a burst of increased activity (contraction) prior to the rise in voice F0 in response to downward pitch-shifted auditory feedback. These same muscles then reduced their activity (relaxation) as the vocal pitch decreased following an upward pitch-shift stimulus. If we assume that the peripheral mechanisms of voice pitch control are relatively identical across musicians and NMs, it can be hypothesized that shorter latency of vocal responses to downward pitch-shifted auditory feedback in AP and RP musicians may be driven by enhanced central nervous mechanisms that speed up the processing of pitch errors for faster behavioral compensations as a result of extensive musical training. Our data support this hypothesis by showing that the amplitude of the P2 ERP components were significantly correlated with the latency of the vocal responses only for downward pitch perturbations in voice auditory feedback (see figure 4). The negative correlation between P2 amplitudes at Fz, FCz and Cz electrodes and vocal response latencies indicated that greater neural activity at these sites was associated with faster reaction times (shorter latencies) to correct for decreasing pitch frequencies (downward shifts) in the auditory feedback. Although the neural processes of faster vocal reactions to feedback changes are not fully understood, the findings of the present study suggest that the P2 responses serve as a robust neural predictor of this facilitated behavioral functionality and can provide us with a quantitative measure to assess vocal controlling abilities in subjects with variable degrees of musical experience. Moreover, due to the fact that the musicians in the present study have received vocal training, it can be argued that in addition to enhanced sensory processing due to musical practice, improved motor functions may also account for better voice pitch control abilities in trained musicians in terms of faster vocal reactions and the ability to re-adjust voice F0 after the removal of the pitch-shift stimulus (see figure 1). This notion is further supported by the results of the musical proficiency tests in this study, showing that the RP and AP musicians scored significantly higher than the NM group in all four tests that evaluated their perceptual (sensory) and vocal production (motor) abilities. However, the results of the present study do not allow us to dissociate the role of sensory-motor functions and therefore, the extent to which these processes may have separately contributed to improved vocal motor controlling abilities can not be fully determined.

Another important finding of the present study is that the right-hemisphere N1 component of ERPs may be a neural marker of enhanced vocal pitch processing in musicians compared with NMs, regardless of whether the musicians possess AP or RP ability. This proposal is supported by our data showing that the N1 responses were significantly stronger in AP and RP musicians compared with the NM group in the right hemisphere but not the left (see figures 2 and 3). This finding is consistent with results from previous studies showing that musicians generate stronger cortical evoked-potentials in response to musical notes and pure tone stimuli (Pantev et al., 1998; A. Shahin, Bosnyak, Trainor, & Roberts, 2003) as well as during performance of a variety of auditory discrimination tasks (Atienza et al., 2003; A. Shahin et al., 2003; Tremblay, Kraus, McGee, Ponton, & Otis, 2001) compared with the NM group. Since the ERPs in the present study were recorded during active vocalization, it is possible that in addition to the primary (Burkard, Don, & Eggermont, 2006) and non-primary (Hari, Aittoniemi, Jarvinen, Katila, & Varpula, 1980) cortical auditory sources, other generators (e.g. motor-related) may have contributed to the N1 responses (for description of other N1 generators see (Naatanen & Picton, 1987)). However, an alternative explanation of the present findings is that the observed N1 differences between the musician

and NM groups may be accounted for by better auditory capability and not necessarily by auditory-motor interactions. These enhanced capabilities may also be related to the increased white matter observed in trained musicians compared with non-musicians (Halwani, Loui, Ruber, & Schlaug, 2011) and between musicians with AP ability compared to those without AP ability (Loui, Li, Hohmann, & Schlaug, 2011). At this point, it is difficult to determine the extent of the relative contribution from purely auditory vs. vocal motor neural sources based on the results of the present study because the ERP responses were not recorded during the a passive listening condition, and therefore, could not be compared with those during vocalization.

More recent studies have also provided evidence that the cerebellum may play a key role in generating predictive feedforward motor commands that drive neural suppression in the auditory modality (Knolle, Schroger, Baess, & Kotz, 2012; Knolle, Schroger, & Kotz, 2012). In these studies, it has been shown that the N1 suppression as a result of listening to self-triggered sounds is significantly reduced in a group of patients with focal cerebellar lesions as compared with a healthy control group. These findings suggest an important dissociation between the neural generators of the N1 and P2 components for the processing of self-generated sounds, indicating that these components may reflect differential cortical and sub-cortical processes that facilitate voice motor control.

With relevance to this discussion, an important question is that of the possible mechanisms in which the N1 and P2 components are involved in vocal pitch error detection and correction. The analysis of these two components in the present study has enabled us to capture the temporo-spatial dynamics of feedback pitch processing during vocal production and control. As discussed earlier, our findings suggest that the N1 potential reflects earlier stages of cortical processing and may be related to detection of the stimulus and is sensitive to its acoustical features such as the direction of the pitch shift in voice auditory feedback. It has also been suggested that the N1 component may reflect improved auditory processing in trained musicians (Pantev, Elbert, Ross, Eulitz, & Terhardt, 1996; Pantev, Hoke, Lutkenhoner, & Lehnertz, 1989), which could possibly result from structural enlargements or extended neural connectivity in the planum temporale as a result of extensive musical practice or innate AP ability (Jancke et al., 2012; Loui et al., 2012; Schlaug et al., 1995).

As for the P2 ERP response, this component is associated with later stages of cortical processing and may reflect functional mechanisms that demand higher level cognitive processes for auditory feedback processing and vocal motor control. It was previously shown that this ERP component is sensitive to feedback pitch error as its amplitude increased as a function of increases in the pitch-shift stimulus magnitude during vocalization (larger feedback error) (Behroozmand et al., 2009). Therefore, based on the findings of the present study, it can be suggested that larger P2 responses to pitch-shifted voice feedback in musicians may relate to training-induced enhancement of neural sensitivity for detecting vocal pitch error in the auditory feedback. This effect may be associated with improved pitch memory processing in trained musicians, which has been suggested to occur in the left hemisphere supramarginal gyrus (Gaab, Gaser, & Schlaug, 2006; Gaab, Gaser, Zaehle, Jancke, & Schlaug, 2003). It is well known that the left hemisphere is dominant for language processing and is more sensitive to rapid temporal fluctuations in fine structures of auditory stimuli (R. J. Zatorre & Belin, 2001). The enhanced left hemisphere P2 response in the AP musicians may reflect processing mechanisms that allow them to utilize language abilities that assign naming conventions to acoustical stimuli as they relate to a musical scale. Moreover, since it is known that the pre-attentive ability to extract linguistic information from complex sounds improves with musical training, AP and RP musicians may develop processing mechanisms reflected in the P2 response that are less well developed in non-musicians (Marie, Magne, & Besson, 2011; A. J. Shahin, 2011). Therefore, another possible

interpretation is that musical training may affect the mechanisms of language processing in such a way as to increase the capacity to more accurately track frequency changes in order to detect vocal pitch errors.

The left-hemisphere difference in P2 responses between AP and RP musicians also raises important questions regarding the functional mechanisms related to this ERP component in trained musicians. Since the behavioral responses did not differ between the AP and RP musicians, it is not likely that the P2 response reflects only motor processes. Thus, the most parsimonious explanation for the P2 difference is that it represents enhanced sensory processing related to AP ability in musicians. Such processing in the AP musicians may reflect a mechanism wherein they are able to precisely identify a musical note (sensory perception) and assign it to an internal representation of vocal pitch (motor representation) within a musical scale. Another possibility is that since the P2 occurs after the beginning of the motor response, the P2 may reflect an evaluation of the error in motor response in the context of their internal representation of the musical scale. The precise details of the neural substrates of this processing are unknown, but as discussed above, Zarate et al. (Zarate et al., 2010b) have proposed a neural network activated by pitch-shifted feedback in singers and non-singers. This network included bilateral planum temporale, posterior superior temporal sulcus, intraparietal sulcus supplementary motor area, anterior cingulate cortex, ventral premotor cortex, anterior insula, right anterior cingulate and putamen. It's difficult to say which of these areas of the left hemisphere may have been activated in the AP musicians of the present study, however, since it is known that the planum temporale is larger in the left hemisphere of AP musicians (Schlaug et al., 1995), it is reasonable to suppose that this area was involved in generating the stronger ERPs as part of the network for detecting voice pitch and the change in direction of the pitch-shifted feedback.

Acknowledgments

This research was supported by NIH Grant No. 1R01DC006243. The authors wish to thank Robert Reinhart, D. Mus., for his help in designing the tests of musical proficiency evaluation.

References

- Atienza M, Cantero JL, Grau C, Gomez C, Dominguez-Marin E, Escera C. Effects of temporal encoding on auditory object formation: a mismatch negativity study. *Brain Res Cogn Brain Res*. 2003; 16(3):359–371. [PubMed: 12706216]
- Bass P, Jacobsen T, Schroger E. Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*. 2008; 70(2):137–143. [PubMed: 18627782]
- Behroozmand R, Karvelis L, Liu H, Larson CR. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clinical Neurophysiology*. 2009; 120(7):1303–1312. [PubMed: 19520602]
- Behroozmand R, Larson CR. Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neurosci*. 2011; 12:54. [PubMed: 21645406]
- Bendixen A, SanMiguel I, Schroger E. Early electrophysiological indicators for predictive processing in audition: a review. *Int J Psychophysiol*. 2012; 83(2):120–131. [PubMed: 21867734]
- Boersma P. Praat, a system for doing phonetics by computer. *Glott International*. 2001; 5(9):341–345.
- Burkard, RF.; Don, M.; Eggermont, JJ. *Auditory Evoked Potentials: Basic Principles and Clinical Application*. 1. Lippincott Williams & Wilkins; 2006.
- Burnett TA, Freedland MB, Larson CR, Hain TC. Voice F0 responses to manipulations in pitch feedback. *J Acoust Soc Am*. 1998; 103(6):3153–3161. [PubMed: 9637026]
- Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc Natl Acad Sci U S A*. 2013

- Chen SH, Liu H, Xu Y, Larson CR. Voice F0 responses to pitch-shifted voice feedback during English speech. *Journal of the Acoustical Society of America*. 2007; 121(2):1157–1163. [PubMed: 17348536]
- Donath TM, Natke U, Kalveram KT. Effects of frequency-shifted auditory feedback on voice F0 contours in syllables. *Journal of the Acoustical Society of America*. 2002; 111(1):357–366. [PubMed: 11831808]
- Eliades SJ, Wang X. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J of Neurophysiology*. 2003; 89(4):2194–2207.
- Eliades SJ, Wang X. Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cerebral Cortex*. 2005; 15(10):1510–1523. [PubMed: 15689521]
- Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*. 2008; 453(7198):1102–1106. [PubMed: 18454135]
- Gaab N, Gaser C, Schlaug G. Improvement-related functional plasticity following pitch memory training. *Neuroimage*. 2006; 31(1):255–263. [PubMed: 16427320]
- Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G. Functional anatomy of pitch memory - an fMRI study with sparse temporal sampling. *Neuroimage*. 2003; 19(4):1417–1426. [PubMed: 12948699]
- Greenlee JD, Behroozmand R, Larson CR, Jackson AW, Chen F, Hansen DR, et al. Sensory-motor interactions for vocal pitch monitoring in non-primary human auditory cortex. *PLoS One*. 2013; 8(4):e60783. [PubMed: 23577157]
- Guenther FH. Cortical interactions underlying the production of speech sounds. *J Commun Disord*. 2006; 39(5):350–365. [PubMed: 16887139]
- Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*. 2006; 96:280–301. [PubMed: 16040108]
- Hackett TA, Stepniewska I, Kaas JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *Journal of Comparative Neurology*. 1998; 394(4):475–495. [PubMed: 9590556]
- Halwani GF, Loui P, Ruber T, Schlaug G. Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Front Psychol*. 2011; 2:156. [PubMed: 21779271]
- Hari R, Aittoniemi K, Jarvinen ML, Katila T, Varpula T. Auditory Evoked Transient and Sustained Magnetic-Fields of the Human-Brain - Localization of Neural Generators. *Experimental Brain Research*. 1980; 40(2):237–240. [PubMed: 7428878]
- Hawco CS, Jones JA, Ferretti TR, Keough D. ERP correlates of online monitoring of auditory feedback during vocalization. *Psychophysiology*. 2009; 46(6):1216–1225. [PubMed: 19674393]
- Heinks-Maldonado TH, Mathalon DH, Gray M, Ford JM. Fine-tuning of auditory cortex during speech production. *Psychophysiology*. 2005; 42(2):180–190. [PubMed: 15787855]
- Heinks-Maldonado TH, Nagarajan SS, Houde JF. Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*. 2006; 17(13):1375–1379. [PubMed: 16932142]
- Hickok G, Houde J, Rong F. Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron*. 2011; 69(3):407–422. [PubMed: 21315253]
- Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience*. 2002; 14(8):1125–1138. [PubMed: 12495520]
- Jancke L, Langer N, Hanggi J. Diminished whole-brain but enhanced peri-sylvian connectivity in absolute pitch musicians. *J Cogn Neurosci*. 2012; 24(6):1447–1461. [PubMed: 22524277]
- Jones JA, Munhall KG. Perceptual calibration of F0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America*. 2000; 108(3):1246–1251. [PubMed: 11008824]
- Kawahara H, Aikawa K. Contributions of auditory feedback frequency components on F0 fluctuations. *Journal of the Acoustical Society of America*. 1996; 100:2825. (A).
- Knolle F, Schroger E, Baess P, Kotz SA. The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *J Cogn Neurosci*. 2012; 24(3):698–706. [PubMed: 22098261]
- Knolle F, Schroger E, Kotz SA. Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*. 2012

- Knolle F, Schroger E, Kotz SA. Prediction errors in self- and externally-generated deviants. *Biol Psychol.* 2013; 92(2):410–416. [PubMed: 23246535]
- Larson CR, Altman KW, Liu H, Hain TC. Interactions between auditory and somatosensory feedback for voice F0 control. *Exp Brain Res.* 2008; 187(4):613–621. [PubMed: 18340440]
- Liu H, Behroozmand R, Bove M, Larson CR. Laryngeal electromyographic responses to perturbations in voice pitch auditory feedback. *J Acoust Soc Am.* 2011; 129(6):3946–3954. [PubMed: 21682416]
- Loui P, Li HC, Hohmann A, Schlaug G. Enhanced cortical connectivity in absolute pitch musicians: a model for local hyperconnectivity. *J Cogn Neurosci.* 2011; 23(4):1015–1026. [PubMed: 20515408]
- Loui P, Zamm A, Schlaug G. Enhanced functional networks in absolute pitch. *Neuroimage.* 2012; 63(2):632–640. [PubMed: 22836173]
- Marie C, Magne C, Besson M. Musicians and the metric structure of words. *J Cogn Neurosci.* 2011; 23(2):294–305. [PubMed: 20044890]
- Müller-Preuss P, Ploog D. Inhibition of auditory cortical neurons during phonation. *Brain Research.* 1981; 215(1–2):61–76. [PubMed: 7260601]
- Naatanen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology.* 1987; 24(4):375–425. [PubMed: 3615753]
- Oostenveld R, Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology.* 2001; 112(4):713–719. [PubMed: 11275545]
- Pantev C, Elbert T, Ross B, Eulitz C, Terhardt E. Binaural fusion and the representation of virtual pitch in the human auditory cortex. *Hear Res.* 1996; 100(1–2):164–170. [PubMed: 8922991]
- Pantev C, Hoke M, Lutkenhoner B, Lehnertz K. Tonotopic organization of the auditory cortex: pitch versus frequency representation. *Science.* 1989; 246(4929):486–488. [PubMed: 2814476]
- Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M. Increased auditory cortical representation in musicians. *Nature.* 1998; 392(6678):811–814. [PubMed: 9572139]
- Parkinson AL, Flagmeier SG, Manes JL, Larson CR, Rogers B, Robin DA. Understanding the neural mechanisms involved in sensory control of voice production. *Neuroimage.* 2012; 61(1):314–322. [PubMed: 22406500]
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The processing of temporal pitch and melody information in auditory cortex. *Neuron.* 2002; 36(4):767–776. [PubMed: 12441063]
- Rauschecker JP. An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hear Res.* 2011; 271(1–2):16–25. [PubMed: 20850511]
- Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience.* 2009; 12(6):718–724.
- Schlaug G, Jancke L, Huang Y, Steinmetz H. In vivo evidence of structural brain asymmetry in musicians. *Science.* 1995; 267(5198):699–701. [PubMed: 7839149]
- Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J Neurosci.* 2003; 23(13):5545–5552. [PubMed: 12843255]
- Shahin AJ. Neurophysiological influence of musical training on speech perception. *Front Psychol.* 2011; 2:126. [PubMed: 21716639]
- Sperry R. Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative & Physiological Psychology.* 1950; 43:482–489. [PubMed: 14794830]
- Takeuchi AH, Hulse SH. Absolute pitch. *Psychol Bull.* 1993; 113(2):345–361. [PubMed: 8451339]
- Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage.* 2008; 39(3):1429–1443. [PubMed: 18035557]
- Tremblay K, Kraus N, McGee T, Ponton C, Otis B. Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear.* 2001; 22(2):79–90. [PubMed: 11324846]
- von Helmholtz, H. *Handbook of physiological optics.* 1. Hamburg: Voss; 1867.
- Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. *Science.* 1995; 269(5232):1880–1882. [PubMed: 7569931]

- Zarate JM, Wood S, Zatorre RJ. Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia*. 2010a; 48:607–618. [PubMed: 19896958]
- Zarate JM, Wood S, Zatorre RJ. Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia*. 2010b; 48(2):607–618. [PubMed: 19896958]
- Zarate JM, Zatorre RJ. Neural substrates governing audiovocal integration for vocal pitch regulation in singing. *Ann N Y Acad Sci*. 2005; 1060:404–408. [PubMed: 16597793]
- Zarate JM, Zatorre RJ. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage*. 2008; 40(4):1871–1887. [PubMed: 18343163]
- Zatorre R, Belin P, Penhune VB. Structure and function of auditory cortex: Music and Speech. *Trends in Cognitive Sciences*. 2002; 6:37–46. [PubMed: 11849614]
- Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. *Cereb Cortex*. 2001; 11(10):946–953. [PubMed: 11549617]
- Zatorre RJ, Perry DW, Beckett CA, Westbury CF, Evans AC. Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proc Natl Acad Sci U S A*. 1998; 95(6):3172–3177. [PubMed: 9501235]

- Left-hemisphere activity is related to enhanced pitch processing in musicians.
- Vocal pitch control is improved in musicians with absolute and relative pitch.
- Musical training leads to faster vocal reactions during voice pitch control.

Figure 1.

The overlaid time course of the behavioral vocal responses to **a)** upward (+100 cents) and **b)** downward (-100 cents) pitch-shift stimuli for non-musicians (NM), relative pitch (RP) musicians and absolute pitch (AP) musicians. The boxplots on the right show the lower quartiles, median and upper quartiles of the peak magnitudes and latencies. The outliers are marked by “○” and the “+” marks in red show the mean values for the magnitude and latency measures. The “*” marks on the vocal response plots show the latencies for which the magnitude of the responses were significantly different for the NMs compared with AP and RP musicians ($p < 0.05$).

Figure 2.

a) The overlaid time course of the ERP responses to upward (+100 cents) pitch-shift stimuli for absolute pitch (AP) musicians, relative pitch (RP) musicians and non-musicians (NM) separately plotted for three different electrode contacts over the left hemisphere (T7), vertex (Cz) and right hemisphere (T8). **b)** Topographical distribution maps of the N1 (140–160 ms) and P2 (240–260 ms) component of the ERPs in response to upward pitch-shift stimuli (+100 cents) for non-musicians (NM), relative-pitch (RP) musicians and absolute-pitch (AP) musicians. **c)** Barplots of the N1 and P2 response amplitudes across three different groups of non-musicians (NM), relative-pitch (RP) and absolute-pitch (AP) musicians on the left and right temporal lobe areas for upward (+100 cents) pitch shift stimuli. The “*” mark in each plot indicates a significant difference ($p < 0.05$).

Figure 3.

a) The overlaid time course of the ERP responses to downward (–100 cents) pitch-shift stimuli for absolute pitch (AP) musicians, relative pitch (RP) musicians and non-musicians (NM) separately plotted for three different electrode contacts over the left hemisphere (T7), vertex (Cz) and right hemisphere (T8). **b)** Topographical distribution maps of the N1 (140–160 ms) and P2 (240–260 ms) component of the ERPs in response to downward pitch-shift stimuli (–100 cents) for non-musicians (NM), relative-pitch (RP) musicians and absolute-pitch (AP) musicians. **c)** Barplots of the N1 and P2 response amplitudes across three different groups of non-musicians (NM), relative-pitch (RP) and absolute-pitch (AP) musicians on the left and right temporal lobe areas for downward (–100 cents) pitch shift stimuli. The “*” mark in each plot indicates a significant difference ($p < 0.05$).

Figure 4.

Results of the correlation analysis using Pearson's method between the amplitude of the P2 responses at Cz electrode (vertex) and the latency of compensatory vocal responses to a) upward (+100 cents) and b) downward (-100 cents) pitch-shift stimuli.

Table 1

The mean (and standard deviation) of non-musician (NM), relative-pitch (RP) and absolute pitch (AP) musicians' scores in four tests of musical proficiency evaluation.

	Non-musician	Relative Pitch Musician	Absolute Pitch Musician
Individual Chromatic Pitch (Part A)	22.35% (13.29)	94.14% (9.17)	99.21% (1.48)
Individual Chromatic Pitch (Part B)	1.25% (0.79)	33.86% (26.01)	96.59% (6.64)
Tonal Sight-Singing	32.5% (37.0)	87.7% (21.9)	97.7% (7.5)
Atonal Sight-Singing	11.25% (18.0)	74.7% (18.4)	100% (0)
Individual Microtonal Pitch	30.6% (30.1)	69.6% (20.9)	86.9% (9.3)
Total Score	21.3% (14.3)	73.0% (8.0)	96.2% (2.3)