Published in final edited form as:

Brain Res. 2012 April 11; 1448: 89–100. doi:10.1016/j.brainres.2012.02.012.

ERP Correlates of Pitch Error Detection in Complex Tone and Voice Auditory Feedback with Missing Fundamental

Roozbeh Behroozmand, Oleg Korzyukov, and Charles R. Larson

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

Abstract

Previous studies have shown that the pitch of a sound is perceived in the absence of its fundamental frequency (F0), suggesting that a distinct mechanism may resolve pitch based on a pattern that exists between harmonic frequencies. The present study investigated whether such a mechanism is active during voice pitch control. ERPs were recorded in response to +200 cents pitch shifts in the auditory feedback of self-vocalizations and complex tones with and without the F0. The absence of the fundamental induced no difference in ERP latencies. However, a right-hemisphere difference was found in the N1 amplitudes with larger responses to complex tones that included the fundamental compared to when it was missing. The P1 and N1 latencies were shorter in the left hemisphere, and the N1 and P2 amplitudes were larger bilaterally for pitch shifts in voice and complex tones compared with pure tones. These findings suggest hemispheric differences in neural encoding of pitch in sounds with missing fundamental. Data from the present study suggest that the right cortical auditory areas, thought to be specialized for spectral processing, may utilize different mechanisms to resolve pitch in sounds with missing fundamental. The left hemisphere seems to perform faster processing to resolve pitch based on the rate of temporal variations in complex sounds compared with pure tones. These effects indicate that the differential neural processing of pitch in the left and right hemispheres may enable the audio-vocal system to detect temporal and spectral variations in the auditory feedback for vocal pitch control.

Keywords

Pitch Processing; Missing Fundamental; Vocal Motor Control; Auditory Feedback; Event-Related Potentials (ERPs)

1. Introduction

The processing of pitch changes in the auditory feedback of self-produced voice has been suggested to play an important role in regulating vocal pitch output during tasks such as speaking or singing. Several studies have demonstrated that humans vocally compensate for shifts in the pitch of their voice auditory feedback by changing their voice fundamental frequency (F0) output in the opposite direction to the pitch shift stimuli (PSS) (Burnett et al., 1998; Chen et al., 2007; Donath et al., 2002; Hain et al., 2000; Kawahara and Aikawa, 1996;

© 2012 Elsevier B.V. All rights reserved.

Corresponding author: Roozbeh Behroozmand, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL, 60208, r-behroozmand@northwestern.edu, Telephone: 847-491-2428, Fax: 847-467-2776.

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.

Xu et al., 2004). This means that an upward shift in voice pitch feedback leads to a decrease in voice F0, and a decrease in pitch feedback leads to an increase in F0. The compensatory nature of vocal responses suggests that the audio-vocal system operates like a negative feedback control system that helps to stabilize voice F0 at an intended frequency. In addition, the short onset latency (~ 100 ms) of these responses suggests they are reflexive in nature and are not controlled by the voluntary mechanisms.

An important step toward understanding the neural mechanisms of audio-vocal control is to understand how voice pitch is processed in the auditory system during vocal production and control. Since a spectrally-rich periodic complex sound such as the human voice consists of multiple frequency components that include voice F0 and its harmonics, it is important to learn whether the auditory system detects pitch changes in voice feedback merely based on the F0, higher harmonic frequencies of the F0, or a combination of all harmonically related frequencies. One possible way to address this question is to study the neural mechanisms of voice pitch control in conditions where the F0 component is either present or missing in the auditory feedback during vocalization. Therefore, in the present study, event-related potentials (ERPs) were recorded from the human brain in response to pitch shift stimuli while the auditory feedback was randomly modified with the subjects' own voice or a computer-generated periodic complex tone in which the F0 was present or missing during vocalization. This method of voice feedback modification also allowed us to investigate whether differences in neural processing of pitch in sounds with and without the F0 may lead to differences in vocal motor control mechanisms that stabilize voice pitch output during vocalization or speaking.

The auditory processing of pitch is central to the perception of melodic, harmonic and prosodic aspects of biologically-significant complex sounds such as speech and music (Trainor, 2008; Zatorre, 2005). Pitch is an auditory perceptual property that allows us to assign periodic complex sounds (e.g. musical notes, human voice, etc.) to relative positions on a frequency-related scale. The neural mechanisms underlying pitch processing have been extensively studied in humans' and animals using electrophysiological and functional imaging techniques. These studies have suggested that two major auditory processing mechanisms are involved in processing and perception of pitch in periodic acoustical stimuli (Bendor and Wang, 2005; Zatorre and Belin, 2001). The first mechanism involves processing the temporal aspects of the stimuli and extracts pitch from the rate of periodic repetitions in the time domain. The second mechanism relies on encoding the spectral cues that exist in the frequency spectrum of periodic complex sounds. Since pitch perception cannot be fully accounted for by the responses of the auditory nerve fibers as a result of peripheral distortion, it has been proposed that pitch processing extends beyond the peripheral auditory system and both the peripheral and central neural mechanisms are involved in resolving pitch based on the spectro-temporal features of complex auditory stimuli (Bendor, 2011).

Although the exact neuronal representation of pitch encoding mechanisms remains unclear, previous studies have provided evidence that brainstem and cortical mechanisms are involved in this process. At the cortical level, the right hemisphere auditory areas are shown to be specialized for processing spectral aspects of complex sounds, whereas the left auditory cortex is more involved in processing temporal aspects of the auditory stimuli (Fujioka et al., 2003; Zatorre and Belin, 2001). These functional hemispheric differences of pitch extraction suggest that speech processing, which generally requires good temporal resolution for rapid encoding of acoustical variations (e.g. rapid sequencing of phonemes, syllables and words), predominantly recruits the left auditory cortex whereas tonal variations that require spectral processing are encoded in the right cortical auditory areas.

A number of studies have demonstrated that humans and animals perceive the F0 of a periodic complex sound even when the actual F0 component is missing (Bendor and Wang, 2005; Tomlinson and Schwarz, 1988; Zatorre, 1988). Although the physical entity of pitch is encoded in the auditory periphery by a frequency-place (tonotopic) representation, the existing behavioral results suggest that the perception of pitch is a unified entity that is not physiologically represented merely by the physical presence of acoustical energy at the F0, but rather is recognized by a pattern that exists in the harmonic frequencies of a complex sound. This characteristic has led to the hypothesis that pitch, like other features of acoustical stimuli such as amplitude, frequency, timing and other parameters, is a basic dimension of the auditory stimulus for which a distinct neural processing mechanism exists in the brain. The developmental changes of pitch processing have been demonstrated in the human brain by showing that 4-months-old infants and older adults neurally respond to the pitch of a missing fundamental complex sound whereas such neural responses were not observed in younger infants (3-months-old) (He and Trainor, 2009). Recent pitch perception theories have proposed that the central pattern recognition processes extracts pitch by combining the temporal and spectral features of complex acoustical stimuli (Cedolin and Delgutte, 2005; Cedolin and Delgutte, 2010), but the underlying neural mechanisms are not well understood.

The present study investigated whether the neural mechanisms of vocal pitch error detection and correction are different in the absence or presence of the fundamental frequency component in the auditory feedback during vocal production. Event-related potentials (ERPs) were recorded in response to pitch shifted auditory feedback when human subjects actively maintained a steady vowel sound /a/ at their conversational pitch and loudness. The voice feedback was randomly modified during vocalization with five types of stimuli: (1) the natural voice signal (Voice), (2) the missing fundamental voice (Voice_{MF}, high pass filtered voice to remove F0), (3) a computer-generated periodic complex tone consisting of one frequency component that tracked and was identical to subject's voice F0 along with its first four harmonics (Tone), (4) a periodic complex tone with missing fundamental (Tone_{MF}), (5) and a synthetic pure sine wave that tracked the voice F0 (F0). The real-time tracking of all stimuli enabled us to generate complex and pure tones that followed natural temporal variations in subjects' voice F0 during vocalization. The objective of this study was to compare ERP latencies and amplitudes to help resolve the question of whether or not there are different neural mechanisms for pitch processing in the auditory feedback when the F0 component is present compared to when it is missing. Therefore, ERPs were compared for Voice vs. Voice_{MF} and Tone vs. Tone_{MF} separately. Furthermore, the ERP responses to Voice and Tone feedback were also separately compared with responses to pitch shifts in a pure sinusoidal tone (F0) in order to investigate whether the neural processing of pitch is different when high-frequency harmonics are present in the frequency spectrum of the signal compared to when they are absent.

This study design allowed us to investigate the underlying neural mechanisms that help detect pitch shifts in the auditory feedback and trigger compensatory vocal responses to correct for vocal pitch errors. The main hypothesis was that there are different types of neural processes of feedback pitch error detection, and therefore there should be different ERPs evoked for pitch perturbations in sounds that include the F0 component compared to those in which the F0 is missing. This hypothesis arises from the fact that previous studies have suggested that different neural mechanisms are involved in resolving pitch in sounds with missing fundamental (Bendor and Wang, 2005; Fujioka et al., 2003; Tomlinson and Schwarz, 1988; Zatorre, 1988). Since the right cortical auditory areas were shown to be specialized for resolving pitch in missing fundamental sounds (Fujioka et al., 2003; Zatorre and Belin, 2001), we predicted that the difference between neural encoding of pitch for feedback with and without the F0 would be more prominently reflected in ERPs in the right

hemisphere. We also hypothesized that the elicited responses to pitch shifts in harmonically-rich sounds (e.g. Voice and complex Tone) would be different from those to shifts in a pure tone that does not include high-frequency harmonics (F0). Lastly, we predicted that the vocal reactions to pitch perturbations would seem nearly identical for sounds with and without the F0 component because the brain is able to extract pitch information even in the absence of the fundamental frequency component of vocalizations.

2. Results

2.1. ERP Responses

Repeated-measures analysis of variances (Rm-ANOVAs) were separately performed on the amplitude and latency of P1, N1, P2 and N2 components from the left-central (C3), vertex (Cz) and right-central (C4) electrodes in order to examine how ERPs are modulated in response to pitch-shifts in periodic complex tones (Tone vs. Tone_{MF}) and voice (Voice vs. Voice_{MF}) feedback with and without the fundamental frequency component. In order to examine laterality, ERP responses to each feedback type were separately analyzed on the left (C3) vs. right (C4) hemisphere electrodes.

Figures 1 and 2 depict the overlaid time courses of the grand averaged (over 14 subjects) ERP responses to pitch shifts in Tone, Tone_{MF}, F0 and Voice, Voice_{MF}, F0 feedback in the left hemisphere (C3), vertex (Cz) and right hemisphere (C4) electrodes separately. The topographical distribution maps of the N1 difference for Tone-Tone_{MF} and Voice-Voice_{MF} are also shown in these figures. In figures 3 and 4, the box plot representation of the latencies and amplitudes of the P1, N1 and P2 components are separately drawn for the left (C3) and right (C4) hemisphere electrodes. The lower, middle and upper lines of each box plot mark the 25th percentiles, median and 75th percentiles of the data and the whiskers extend to the most extreme data points that are not considered outliers, and each outlier is plotted as an individual point. The “+” marks in each box plot indicates the mean value of the data.

2.1.1. ERP Latencies—The latencies of the P1 responses were significantly shorter ($F(1,13)=4.74$, $p=0.048$) in response to pitch shifts in Tone compared with F0 only in the left hemisphere (figures 1a and 3a). A laterality main effect was found for the P1 latencies to be significantly shorter ($F(1,13)=6.14$, $p=0.028$) in the left compared with the right for pitch shifts in Tone feedback (figure 3a). No significant laterality main effect was found for P1 latencies in response to pitch shifts in Voice, Voice_{MF}, Tone_{MF} and F0 in the left compared with the right side.

Results of the analysis showed that the latencies of the N1 responses to pitch shifts in Tone ($F(1,13)=19.19$, $p=0.001$) and Tone_{MF} ($F(1,13)=35.47$, $p=0.000$) were significantly shorter than those for pitch shift in F0 feedback in the left hemisphere (figures 1a and 3b). No significant laterality main effect was found for N1 latencies in response to pitch shifts in Tone, Tone_{MF} and F0 in the left compared with the right side. In comparison with F0 feedback, N1 responses had significantly shorter latencies for Voice ($F(1,13)=14.7$, $p=0.002$) and Voice_{MF} ($F(1,13)=48.33$, $p=0.000$) feedback in the left hemisphere. No significant laterality main effect was found for N1 latencies in response to pitch shifts in Voice, Voice_{MF} and F0 on the left compared with the right side.

Results of the analysis for P2 and N2 latencies did not show any significant main effects of feedback or laterality for pitch shifts in periodic complex tone or voice auditory feedback.

2.1.2. ERP Amplitudes—Results of the analysis of P1 amplitudes did not show any significant main effects of feedback or laterality for pitch shifts in periodic complex tone or voice auditory feedback (figures 1a and 4a).

For the N1 component, the pitch shifts in periodic complex tones elicited responses that had significantly larger amplitudes ($F(1,13)=5.19$, $p=0.04$) for Tone compared with Tone_{MF} feedback only on the right hemisphere (see figures 1a and 4b). This effect is illustrated by a right-hemispheric negativity in the topographical distribution maps of the difference in N1 amplitudes for Tone vs. Tone_{MF} feedback (Tone – Tone_{MF}) in figure 1b. In comparison with pure tones (F0), the N1 amplitudes were significantly larger for Tone feedback in both hemispheres (Left: $F(1,13)=18.07$, $p=0.001$, Right: $F(1,13)=7.42$, $p=0.017$) but for Tone_{MF}, the N1 amplitude was larger only in the left hemisphere ($F(1,13)=9.24$, $p=0.009$). No laterality effect was found for N1 amplitudes on the left vs. right side for Tone, Tone_{MF} or F0 feedback. The pitch shifts in Voice and Voice_{MF} feedback elicited N1 responses that had significantly larger amplitudes (Voice-Left: $F(1,13)=10.25$, $p=0.014$, Voice-Right: $F(1,13)=11.01$, $p=0.008$ and Voice_{MF}-Left: $F(1,13)=11.39$, $p=0.005$, Voice_{MF}-Right: $F(1,13)=13.33$, $p=0.001$) compared with F0 on both hemispheres (see figures 2a and 4b). As can be seen in figure 2b, the distribution of N1 differences for Voice vs. Voice_{MF} does not show any hemispheric dominance. In addition, no laterality effect was found for N1 amplitudes on the left vs. right side for Voice, Voice_{MF} or F0 feedback.

For the P2 ERP components, pitch shifts elicited responses that had significantly greater amplitudes in both hemispheres for Tone (Left: $F(1,13)=14.09$, $p=0.002$, Right: $F(1,13)=18.95$, $p=0.001$) and Tone_{MF} (Left: $F(1,13)=13.72$, $p=0.003$, Right: $F(1,13)=18.88$, $p=0.001$) compared with F0 feedback (see figures 1a and 4c). No main effect of laterality was found in P2 response amplitudes on the left vs. right for Tone, Tone_{MF} and F0 feedback. The P2 responses to pitch shifts in Voice (Left: $F(1,13)=33.7$, $p=0.000$, Right: $F(1,13)=9.54$, $p=0.009$) and Voice_{MF} feedback (Left: $F(1,13)=19.35$, $p=0.001$, Right: $F(1,13)=16.22$, $p=0.002$) also had significantly larger amplitudes compared with F0 in both hemispheres (see figures 2a and 4c). No significant main effect of laterality was found for P2 amplitudes for left vs. right side in response to pitch shifts in Voice, Voice_{MF} and F0 feedback.

The analysis of the N2 (negativity around 300 ms) peak amplitudes did not reveal any main effect of feedback or laterality for Tone, Tone_{MF} and F0. However, results indicated larger N2 responses (more negative) to pitch shifts in both hemispheres for Voice_{MF} compared with Voice (Left: $F(1,13) = 5.94$, $p=0.03$, Right: $F(1,13) = 8.53$, $p=0.012$) and F0 (Left: $F(1,13) = 11.29$, $p=0.005$, Right: $F(1,13) = 8.67$, $p=0.011$) feedback (figure 2a).

2.2. Vocal Responses

All of the subjects responded to the upward pitch-shift stimuli with a compensating, downward vocal response that lowered their voice F0. Figure 5 displays the grand averaged voice F0 responses to +200 cents pitch shifts in the five different feedback types (Voice, Voice_{MF}, Tone, Tone_{MF} and F0) tested in this study. It is apparent from this figure that the response peak magnitudes were largest for pitch shifts in Voice feedback (mean: -5.29 cents) compared with other feedback types. However, no evident difference can be seen among vocal response peak magnitudes for Voice_{MF} (-3.62 cents), Tone (-3.71 cents), Tone_{MF} (-3.57 cents) and F0 (-3.70 cents) feedback. The latencies of the peak of vocal responses were about the same for all feedback types and their mean values were 249, 241, 257, 253 and 258 ms for Voice, Voice_{MF}, Tone, Tone_{MF} and F0, respectively. A one-way (5) Rm-ANOVA on the peak magnitude of vocal responses with factors including feedback (Voice, Voice_{MF}, Tone, Tone_{MF} and F0) revealed a significant main effect of feedback ($F(4,52)=5.20$, $p=0.001$). Post-hoc tests revealed that the peak magnitude of vocal responses

to pitch shifts in Voice feedback were significantly larger (more negative) compared with Voice_{MF} ($p=0.000$), Tone ($p=0.000$), Tone_{MF} ($p=0.007$) and F0 ($p=0.001$) feedback (figure 5). However, no significant difference was revealed between vocal responses to pitch shifts in Voice_{MF}, Tone, Tone_{MF} and F0 feedback. A one-way (5) Rm-ANOVA on the latency of vocal response peaks did not reveal any significant main effects.

3. Discussion

3.1. ERP Latencies

The present study investigated the ERP correlates of pitch shift detection in the auditory feedback of naturally-produced human vocalizations and artificially-generated periodic complex sounds in which the F0 component was either present or missing. Results of the analysis for ERP latencies showed a laterality effect for the P1 latencies being significantly shorter in response to pitch shifts in complex tone (Tone) feedback on the left compared with right hemisphere, indicating a faster processing speed for detecting pitch changes in the left cortical auditory areas (figure 3a). Previous evidence for left hemispheric specialization for temporal and right hemisphere for spectral processing (Fujioka et al., 2003; Zatorre and Belin, 2001) suggests that decoding frequency by tracking the rate of periodic repetition (temporal features) in the left cortical auditory areas may result in faster neural processing compared with spectral processing in the right hemisphere. However, our results indicated that the P1 latencies did not reflect such a laterality effect for Voice or pure tone (F0) feedback. The failure to see a significant difference in the P1 latency in the left compared with right hemisphere for the F0 signal (figure 3a), suggests that the speed of pitch processing for a stimulus with a single frequency component (F0) based on temporal (in the left) and spectral (in the right) variations is nearly identical. On the other hand, the shorter P1 latencies found for the complex Tone compared with F0 in the left hemisphere, suggest that the speed of auditory neural processing for a stimulus with multiple frequency components (periodic complex tone) may be faster in the left hemisphere as compared with a pure sinusoidal tone. It is unlikely that this effect could be due to basilar membrane conduction velocities for complex vs. pure tone feedback since this would only account for a fraction of a millisecond of the latency difference in the elicited ERPs. One possibility for the shorter latency for the complex tones that should be considered is that they would excite a larger population of auditory nerve fibers along the basilar membrane, and consequently, the temporal integration of this peripheral neural activity will result in the activity of a greater number of higher-level (e.g. brainstem or cortical) neurons. The larger number of excited neurons might result in faster pitch encoding for complex compared with pure tones that excites a smaller number of neurons.

The advantage of the left auditory cortical areas in faster processing of pitch variations in complex stimuli is more obvious in the latencies of the N1 ERP components. Our results showed that the N1 had a significantly shorter latency in response to Tone and Tone_{MF} compared with F0 in the left hemisphere (figures 1a and 3b). The N1 ERP also had a shorter latency in the left hemisphere for the Voice and Voice_{MF} compared with the F0 feedback (figures 2a and 3b). However, no significant N1 latency difference was found in the left side between Tone vs. Tone_{MF} and Voice vs. Voice_{MF} feedback, suggesting that in comparison with complex sounds that include the F0, resolving pitch in sounds with missing fundamental may involve mechanisms that have nearly the same speed of processing.

The analysis of the P2 component latencies did not reveal any trace of left hemisphere dominance in processing speed of pitch variations in periodic complex sounds (e.g. voice or complex tones) compared with F0 feedback. Although the primary neural generators of the P1 and N1 have been identified within the primary (Burkard et al., 2006) and secondary (Hari et al., 1980) cortical auditory areas (for description of other P1 and N1 generators see:

(Naatanen and Picton, 1987)), little is known about the generators of the P2 component. The P2 component may arise from distributed sources that include auditory and vocal motor related areas that are active during or after vocal production and control. A study has suggested that the auditory generators of the P2 component are different from those of the N1 in terms of their locations and functional roles as physiological markers of perceptual auditory training, learning and memory during passive listening to repetitive presentation of sounds (Ross and Tremblay, 2009). The neural generators of the P2 were found to be more anterior to those for the N1 within the primary auditory cortex and the changes in P2 amplitude have been suggested to reflect mechanisms that involve stimulus feature evaluation, plastic neural reorganization and performance consolidation in an auditory training task. In contrast, changes in the N1 responses during training sessions were shown to occur in a different time-scale as compared with P2 and the N1 suppression was suggested to indicate short-term habituation to repetitive sounds that prevent the brain from responding to stimuli without new information.

In addition, a few recent studies have shown that the P2 component is sensitive to the degree of pitch error in voice auditory feedback (Behroozmand et al., 2009; Liu et al., 2010a), harmonic complexity (Behroozmand et al., 2011) and the direction of pitch shift stimuli (Liu et al., 2010b). Our findings suggest that, in contrast with P1 and N1 components, the P2 responses in the present study do not seem to serve as a neural indicator of the speed of processing for detecting pitch changes in different feedback types such as voice, complex or pure tone stimuli. In fact, the exact functions of the neural generators giving rise to the P2 ERP component remain to be elucidated.

In general, the analysis of the ERP latencies in the present study suggest that, regardless of the presence or absence of the fundamental, pitch variations in periodic complex sounds are resolved faster compared with pure tones in the earlier stages of the neural processing (indexed by P1 and N1 components) in the left hemisphere. These findings are in contrast with findings of an earlier study by Fujioka et al. (Fujioka et al., 2003) in which it was shown that the latencies of the N1m responses were delayed in response to the onset of missing fundamental complex sounds compared with pure tones. Several important distinctions between the experimental setup of the present study and that by Fujioka et al. (Fujioka et al., 2003) may explain the contradiction in timing of the neural responses. First, N1m responses in Fujioka et al.'s (Fujioka et al., 2003) study were elicited in response to the onset of pure tones or of complex tones with missing fundamental, and therefore, may not only code for pitch, but they could also have represented other stimulus parameters such as onset time, loudness etc. In contrast, the N1 responses in the present study were elicited in response to a pitch shift stimulus (not sound onset) within a continuous voice, complex or pure tone and reflected cortical neural mechanisms of encoding pitch variation in the auditory feedback in the presence or absence of the F0.

Therefore, it seems reasonable to suggest that the cortical generators of the N1m component in Fujioka et al.'s (Fujioka et al., 2003) study are different from those for N1 in the present study. Second, Fujioka et al. (Fujioka et al., 2003) did not test for complex tones that included the fundamental (e.g. Voice and Tone in this study), and it is not possible to determine, based on their findings, whether the observed latency effect is directly related to encoding pitch rather than acoustical complexity in the tested periodic complex stimuli compared with pure tones. Third, the ERPs in the present study were elicited in response to pitch-shifted feedback during vocalization of a steady vowel sound, and it is possible that some vocal motor-related components may have contributed to faster processing of the auditory feedback information in missing fundamental compared with pure tones. Evidence from previous studies in humans (Behroozmand and Larson, 2011; Gunji et al., 2001; Heinks-Maldonado et al., 2005; Houde et al., 2002) and non-human primates (Eliades and

Wang, 2003; Eliades and Wang, 2005) have suggested that the vocal motor mechanisms may be involved in modulating the neural processing of auditory feedback during vocal production. These studies have demonstrated that the auditory responses to self-produced vocalizations are suppressed during active vocal production compared with passive listening to the playback, and also have shown that the vocalization-induced suppression begins prior to the onset of vocalizations. A more recent study (Eliades and Wang, 2008) have suggested that the suppressed cortical auditory neurons are more responsive to alterations in voice feedback pitch, suggesting that the efference projections from the vocal motor to auditory system may change tuning properties of cortical auditory neurons in such a way as to increase their sensitivity to detect pitch perturbations in voice feedback. Results of the present study regarding the left-hemisphere dominance in performing faster processing suggests that similar vocal motor mechanisms (e.g. efference system) may as well be involved to enhance the speed of neural processing in cortical auditory neurons for detecting pitch changes in voice feedback.

Another possibility is that the magnetic N1m component may not share a common neural generator with its electrical counterpart (N1) for identical auditory stimuli. As discussed by Fujioka et al. (Fujioka et al., 2003), the response latency effect for magnetic (MEG) N1m in their study was not consistent with results of other studies that investigated the latency effect in a similar paradigm using the electrical N1 component (Crottaz-Herbette and Ragot, 2000; Ragot and Lepaul-Ercole, 1996).

3.2. ERP Amplitudes

The analysis of the ERP amplitudes indicated differences in the neural processing of pitch shifts in the feedback of periodic complex tones with and without the F0 component (Tone vs. Tone_{MF}). The N1 ERP amplitude was significantly larger for Tone compared with Tone_{MF} only over the right hemisphere (figures 1a and 4b). This hemispheric lateralization is more clearly illustrated in the topographical representation of N1 difference maps spanning the right temporal lobe for the Tone vs. Tone_{MF} feedback in figure 1b. This finding suggests that, in comparison with a complex tone stimulus that includes the fundamental, the pitch of a periodic complex tone with missing fundamental is possibly resolved by a different cortical auditory mechanism in the right hemisphere.

The notion of a different pitch processing mechanism for sounds with missing fundamental is supported by behavioral evidence in humans and animals demonstrating that the presence of the F0 component is not a necessary factor for pitch perception in periodic complex auditory stimuli (Bendor and Wang, 2005; Tomlinson and Schwarz, 1988; Zatorre, 1988). This phenomenon has been suggested to be accounted for by a central pattern recognition mechanism that enables the auditory system to resolve pitch based on a pattern that exists between the harmonic frequencies of complex tone stimuli (Winkler et al., 1997). Lesions to the right but not left primary auditory cortical areas (specifically Heschl's gyrus) have been shown to impair the ability to perceive pitch variations in complex tones with missing fundamental, but when the fundamental was included in a control condition, the pitch perception ability remained intact (Zatorre, 1988). This latter piece of evidence suggests that, in addition to spectral processing, the right hemisphere is more specifically specialized to resolve pitch when the fundamental frequency is missing. Source localization of the neural generators of the magnetoencephalographic (MEG) counterpart of the electrical N1 component (N1m) has led to the identification of a tonotopically organized representation of sounds with missing fundamental that was more pronounced in the right-hemisphere cortical auditory areas, with the reconstructed dipoles moving more posteriorly as the sound frequency increased (Fujioka et al., 2003). These findings further support the notion of right hemisphere dominance for encoding pitch in periodic complex sounds based on their spectral contents.

Due to the fact that the pitch shift stimulus induces variations in the temporal (repetition rate) and spectral (harmonic shifts) features of the auditory feedback, the elicited ERPs in the present study reflect the neural mechanisms that encode pitch based on changes in time and frequency domains. However, our finding regarding the right-hemisphere difference in N1 amplitudes for Tone vs. Tone_{MF} (figures 1a, 1b and 4b) suggest that detecting pitch changes in the auditory feedback may predominantly rely on the spectral rather than temporal features. The absence of a significant N1 difference in the left hemisphere for Tone vs. Tone_{MF} feedback suggests that resolving pitch changes based on the temporal characteristics of stimuli in the left cortical auditory areas may be driven by similar mechanisms for complex tones with and without the fundamental frequency.

However, In contrast with artificially-generated periodic complex feedback, no difference was found between the amplitudes of the N1 responses to Voice vs. Voice_{MF} feedback in the right hemisphere, suggesting that there may be a difference between the mechanisms for resolving pitch in a naturally produced human vocalization with missing fundamental and for a complex tone with missing fundamental. This difference may be associated with a more complex harmonic structure of naturally-produced voice that can be used for pitch extraction. It is likely that the brain resolves the pitch in missing fundamental voice feedback by processing high frequency spectral components that were absent in the periodic complex tones tested in the present study. In addition, the perceptual grouping of harmonics by means of spectral cues conveyed via formant frequencies of specific vowel sounds may also lead to a different pitch extraction mechanism as compared to artificial periodic complex tones. However, this is a difficult conclusion to accept if there is equal complexity in the natural vocal and an artificial complex tone. An alternative possibility is that with natural self-generated vocalizations, there are other sources of feedback, such as the bone conducted feedback source or proprioceptive/kinesthetic feedback pathways which may impact the perceptual process.

With the exception to the amplitude of the N1 responses for Tone_{MF} feedback in the right hemisphere, the amplitude of the N1 and P2 responses were significantly larger (N1 with greater negativity and P2 with greater positivity) in response to Voice, Voice_{MF}, Tone and Tone_{MF} compared with F0 in both left and right hemispheres (see figures 1a, 2a, 4b and 4c). This finding suggests that the neural generators of the N1 and P2 components may be more responsive to pitch changes that occur in periodic complex sounds as compared with a pure sinusoidal tone. This notion is supported by findings of an earlier study suggesting that one of the functions of the N1 and P2 generators may be to code for the extent of acoustical complexity in auditory stimuli (Behroozmand et al., 2011). However, the only exception was that there was no significant difference between N1 amplitudes for Tone_{MF} and F0 feedback in the right hemisphere, but instead, the N1 amplitude in the right hemisphere was significantly larger for Tone compared with Tone_{MF} feedback. As discussed earlier, this finding may indicate that in addition to encoding acoustical complexity, the neural generators of the N1 in the right cortical areas may also be involved in resolving the pitch in periodic complex sound with missing fundamental. This suggestion is supported by findings of the previous studies demonstrating that the N1 generators activated by the missing fundamental periodic complex stimuli are located within tonotopically-organized cortical auditory areas that are activated in response to pure sinusoidal tones (Fujioka et al., 2003; Pantev et al., 1989; Pantev et al., 1996). A recent study in non-human primates (Bendor and Wang, 2005) has identified pitch-sensitive neuronal units in a restricted low-frequency region near the antero-lateral border of the primary auditory cortex that respond to the spectral pitch of pure sinusoidal tones, and at the same time, are driven by periodic complex tones with missing fundamental. Interestingly, these pitch-sensitive neurons were shown to be non-responsive to complex sounds that included the fundamental, suggesting that they might be a part of the neural mechanisms that resolves pitch rather than the frequency

components. The findings of the neuroimaging studies have suggested a consistent cortical auditory area for pitch-sensitive neurons in humans (Patterson et al., 2002; Penagos et al., 2004).

Moreover, the analysis of the N2 component (negativity around 300 ms) showed bilaterally larger responses to pitch shifts in Voice_{MF} compared with Voice and F0 feedback (figure 2a). One challenge in interpreting the observed effect on the amplitudes of the N2 in the present study is that the latency of this component (~300 ms) is longer than the PSS duration (200 ms) and therefore, it is likely that the N2 reflects overlapping responses to PSS onset and offset simultaneously. For example, if the PSS offset elicits an N1 response, this response would superimpose the actual N2 response elicited by the onset of the PSS. This problem would make it difficult to discuss the functional role of the N2 generators for processing pitch in missing fundamental voice feedback. Despite the fact that our results may not be convincing enough to draw strong conclusions about the role of N2 generators for encoding pitch, it is likely that pitch processing in sounds with missing fundamental frequency may be driven by different neural mechanisms at higher stages of cortical processing. Although previous studies have provided evidence that the generators of the N2 component in the parietal cortex may reflect higher level cognitive functions such as memory, attention, emotion and semantic processing (Carrette et al., 1997; Franklin et al., 2007), the role of this component in voice feedback pitch processing remains to be elucidated.

3.3. Vocal Responses

Although the analysis of the P1 and N1 latencies have indicated faster neural processing of pitch shifts in the left hemisphere for complex sounds (e.g. voice or complex tones with and without the F0) compared with pure tones feedback, vocal responses to pitch-shifts in all tested feedback types including Voice, Voice_{MF}, Tone, Tone_{MF} and F0 seem to reach their peaks at about the same time (figure 5). This effect suggests that faster neural processing of auditory feedback during vocalization may not necessarily lead to a faster vocal reaction for stabilizing vocal pitch output. One possible explanation for this effect may be that the vocal responses are likely to be more variable across different subjects due to the fact that they can be affected by neural mechanisms in different areas (e.g. cortical, cerebellum, limbic system, brainstem) and varying biomechanical and muscular parameters of the laryngeal and respiratory systems (Davis et al., 1993; Jurgens, 2008; Larson, 1991; Ludlow, 2005; Titze, 1994). The fact that no significant difference was found between the latencies of the vocal responses to Voice, Voice_{MF}, Tone and Tone_{MF} compared with F0 indicates that faster left-hemispheric cortical processing of pitch shifts, as indexed by shorter latencies of the P1 and N1 ERP components, may not directly translate into faster vocal reactions to feedback pitch perturbations.

However, in terms of the magnitude of vocal responses to pitch-shifted feedback, our results showed that subjects compensated with a significantly larger vocal response magnitude for pitch shifts in Voice compared with other feedback types such as Voice_{MF}, Tone, Tone_{MF} and F0 feedback (figure 5). This observation supports a previous finding (Sivasankar et al., 2005) and contradicts our primary hypothesis about similar audio-vocal mechanisms for pitch compensation for complex sounds (e.g. human voice) with and without the fundamental. Although the absence of significant differences between the latencies and amplitudes of ERPs for Voice and Voice_{MF} may suggest similar neural processing of pitch variations for these two types of feedback, vocal compensations seem to be more effective for pitch changes in the feedback of natural voices that include the fundamental frequency component. This effect indicates that pitch variations in naturally-produced human vocalization are more effectively processed by the audio-vocal mechanisms that underlie pitch error detection and correction during vocal production or speaking. These observations

also strongly suggest that the ERPs reported in the present study probably do not relate to motor mechanisms of voice control, i.e., they are mainly driven by auditory sensory mechanisms.

In summary, the present study provided supporting evidence for the existence of a different neural mechanism for the processing of pitch perturbations in the auditory feedback of periodic complex sounds with and without the fundamental. Our results suggest hemispheric specialization of the right cortical auditory areas for encoding pitch in periodic complex tones with missing fundamental, as indexed by the difference in the N1 amplitudes in the right hemisphere for Tone vs. Tone_{MF} feedback. In addition, our results suggested the P1 and N1 components serve as neural indicators of faster processing for pitch perturbations in complex auditory stimuli compared with pure tones. Consequently, pitch shifts in complex sounds were shown to elicit N1 and P2 ERP responses that had larger peak amplitudes as compared with pure tone feedback. However, the shorter latencies of the P1 and N1 components along with larger amplitudes of the N1 and P2 components did not seem to directly translate into shorter latencies or larger magnitudes of the compensatory vocal reactions in response to pitch-shifted auditory feedback. The analysis of vocal responses showed a significantly larger response magnitude only for naturally-produced human vocalization (Voice) as compared with other types of tested feedback stimuli. This effect suggests that the underlying neural mechanisms of audio-vocal integration may process pitch variations more effectively in voice feedback. In general, our findings indicate that the hemispheric specialization of left and right cortical auditory areas for processing temporal and spectral variations in voice feedback enables the audio-vocal system to detect pitch errors and perform vocal motor control during vocalization or speaking.

4. Experimental Procedures

4.1. Subjects

Fourteen right-handed native speakers of American English (9 females and 5 males, 19–28 years of age, mean=22.6 years) participated in this study. All subjects passed a bilateral pure-tone hearing screening test at 20 dB sound pressure level (SPL) (octave frequencies between 250–8000 Hz) and reported no history of neurological disorders, voice or musical training. All study procedures, including recruitment, data acquisition and informed consent were approved by the Northwestern University institutional review board, and subjects were monetarily compensated for their participation.

4.2. Experimental design

The experiment consisted of three blocks of trials of active vocalization. During each block, subjects were asked to sustain the vowel sound /a/ for approximately 2–3 seconds at their conversational pitch and loudness. This vocal task was repeated 150 times during each block while subjects took short breaks (1–2 seconds) between successive utterances. During each vocalization trial, subjects were presented with one of the randomly chosen types of auditory feedback, including their own voice (Voice), their own voice with missing fundamental (Voice_{MF}), a periodic complex Tone that consisted of sinusoids with F0 and its first, second, third and fourth harmonic frequencies (Tone: F0+H1+H2+H3+H4), a periodic complex tone with missing fundamental (Tone_{MF}: Tone-F0 or H1+H2+H3+H4) and a pure sinusoidal tone at the fundamental frequency of their own voice (sine wave at F0). Figure 6 shows the experimental setup, and the test sounds along with their frequency spectra are presented in figure 7. During each vocalization trial, the randomly chosen voice, periodic complex or pure tone feedback was pitch shifted one time in the upward direction at +200 cents magnitude for 200 ms, with a randomized stimulus onset time between 500–1000 ms after vocal onset (figure 6). For the periodic complex tone feedback, the F0 and its harmonics had

equal amplitudes and the relative phase difference between them was zero. The intensities of voice, periodic complex and pure tone feedback were calibrated at about 80 dB SPL. The total duration of each block of trials was approximately 10–15 minutes.

4.3. Instrumentation

Subjects were seated in a sound-treated room in which their voice was picked up with an AKG boomset microphone (model C420). A Mackie mixer (model 1202-VLZ3) was used to amplify and split the subjects' voice signal between three separate channels (figure 6). The second channel was fed to a Kay Visipitch module (Kay Elemetrics, Model 6095/6097) to track voice F0 online by generating a DC signal that was proportional to voice F0 (analog F0). We tested the analog F0 signal and confirmed that the Kay Visipitch output was linear (6 mV/Hz) over the range of human voice F0 (75–500 Hz). The analog F0 signal was read by the Max/Msp program (v.5.0, Cycling 74) and used to control the output frequency of five oscillators through voltage-controlled gates. Oscillators one to five were set to generate sinusoidal tones at F0, 2×F0 (H1), 3×F0 (H2), 4×F0 (H3) and 5×F0 (H4) frequencies, respectively. For generating the missing fundamental voice and periodic complex tone feedback, the fundamental frequency was removed from the voice (Mackie's first channel) and periodic complex tones (oscillator outputs) using a high-pass filter (cutoff: 1.5×F0, slope: −24dB/oct) embedded in the Eventide Eclipse Harmonizer module. Prior to the onset of vocalizations, a random generator function was used to generate a random number between 1 and 5. The Max/Msp program modified the subject's auditory feedback with Voice, Voice_{MF}, Tone, Tone_{MF} and F0 (pure tone) corresponding to the pre-selected random numbers 1 to 5, respectively. The Max/Msp routine also used these randomly generated numbers to adjust the output gain of the Eventide harmonizer so that feedback loudness was nearly the same for all stimuli. At conversational F0, the subjects maintained their vocalizations at a loudness of about 70 dB SPL, and the feedback (Voice, Voice_{MF}, Tone, Tone_{MF} and F0) was delivered through Etymotic earphones (model ER1-14A) at about 80 dB SPL. The 10 dB gain between voice and feedback channels was used to partially mask air-born and bone-conducted voice feedback. 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.

The voice signal from the Mackie's third output channel was fed to a voice onset detector (VOD) module implemented in Max/Msp to detect vocal onset (figure 6). At the onset of each vocalization, one of the five randomly chosen feedback signals (Voice, Voice_{MF}, Tone, Tone_{MF} and F0) was delivered to the subject, and one pitch shift stimulus (200 cents magnitude and 200 ms duration) was delivered to the corresponding auditory feedback using the Eventide Eclipse Harmonizer. All parameters of the pitch shift stimulus such as time delay from vocal onset, inter-stimulus intervals (ISI), duration, direction and magnitude were controlled by the Max/MSP software. The Max/Msp software also generated a TTL (transistor-transistor logic) pulse to mark the onset of each PSS for synchronized averaging of the recorded brain activity and vocal responses. 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).

4.4. 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 and 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.

The BrainVision Analyzer 2 software was used to analyze recorded EEG signals in order to calculate ERPs in response to PSS for Voice, Voice_{MF}, Tone, Tone_{MF} and F0 feedback. 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 PSS. Following segmentation, artifact rejection was carried out by excluding epochs with EEG or EOG amplitudes exceeding ± 50 μ 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. A minimum number of 100 epochs was averaged for each condition. The data were grand averaged over 14 subjects for each condition separately. For each subject, the latency and amplitude of the P1, N1 and P2 ERP components were extracted by finding the most prominent negative and positive peaks in 50 ms-long time windows centered at 70, 120, 210 and 300 ms, respectively. These time windows were selected based upon visual inspection of the grand averaged responses.

In order to ensure that the extracted ERP components had amplitudes above the noise level, their mean amplitudes were compared against zero using a two-tailed one-sample t-test. Results indicated that the mean of P1, N1, P2 and N2 peak amplitudes for all conditions on the left hemisphere (C3), vertex (Cz) and right hemisphere (C4) electrodes were significantly different than zero ($p < 0.05$). The lowest t-score value was found for the P1 amplitude in response to pitch shifts in pure tone (F0) feedback in the left hemisphere C3 electrode ($t(13) = 5.29$, $p = 0.000$).

4.5. 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 then converted from Hertz to Cents scale using the formula: $\text{Cents} = 1200 \times \log_2(F_2/F_1)$ in which F_1 and F_2 are the pre-stimulus and post-stimulus pitch frequencies. The extracted pitch contours were segmented into epochs ranging from –100 ms pre- and +500 ms post-stimulus time intervals and averaged separately for different feedback complexities in each subject. The magnitude and latency of vocal responses were calculated by finding the most prominent peak in 150 ms-long time windows centered at 250 ms.

Acknowledgments

This research was supported by a grant from NIH, Grant No. 1R01DC006243.

References

- Behroozmand R, Karvelis L, Liu H, Larson CR. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clinical Neurophysiology*. 2009; 120:1303–12. [PubMed: 19520602]
- Behroozmand R, Korzyukov O, Larson CR. Effects of voice harmonic complexity on ERP responses to pitch-shifted auditory feedback. *Clinical Neurophysiology*. 2011

- Behroozmand R, Larson CR. Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*. 2011; 12:54. [PubMed: 21645406]
- Bendor D, Wang X. The neuronal representation of pitch in primate auditory cortex. *Nature*. 2005; 436:1161–5. [PubMed: 16121182]
- Bendor D. Understanding how neural circuits measure pitch. *J Neurosci*. 2011; 31:3141–2. [PubMed: 21368024]
- Boersma P. Praat, a system for doing phonetics by computer. *Glott International*. 2001; 5:341–345.
- Burkard, RF.; Don, M.; Eggermont, JJ. Auditory Evoked Potentials: Basic Principles and Clinical Application. Lippincott Williams & Wilkins; 2006.
- Burnett TA, Freedland MB, Larson CR, Hain TC. Voice F0 Responses to Manipulations in Pitch Feedback. *Journal of the Acoustical Society of America*. 1998; 103:3153–3161. [PubMed: 9637026]
- Carretie L, Iglesias J, Garcia T, Ballesteros M. N300, P300 and the emotional processing of visual stimuli. *Electroencephalogr Clin Neurophysiol*. 1997; 103:298–303. [PubMed: 9277632]
- Cedolin L, Delgutte B. Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve. *J Neurophysiol*. 2005; 94:347–62. [PubMed: 15788522]
- Cedolin L, Delgutte B. Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve. *J Neurosci*. 2010; 30:12712–24. [PubMed: 20861376]
- 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; 121:1157–1163. [PubMed: 17348536]
- Crottaz-Herbette S, Ragot R. Perception of complex sounds: N1 latency codes pitch and topography codes spectra. *Clin Neurophysiol*. 2000; 111:1759–66. [PubMed: 11018489]
- Davis, PJ.; Bartlett, DJ.; Luschei, ES. Coordination of the respiratory and laryngeal systems in breathing and vocalization. In: Titze, IR., editor. *Vocal Fold Physiology*. *Frontiers in Basic Science*. Singular; San Diego: 1993. p. 189-226.
- 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; 111:357–366. [PubMed: 11831808]
- Eliades SJ, Wang X. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J of Neurophysiology*. 2003; 89:2194–207.
- Eliades SJ, Wang X. Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cerebral Cortex*. 2005; 15:1510–23. [PubMed: 15689521]
- Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*. 2008; 453:1102–6. [PubMed: 18454135]
- Franklin MS, Dien J, Neely JH, Huber E, Waterson LD. Semantic priming modulates the N400, N300, and N400RP. *Clin Neurophysiol*. 2007; 118:1053–68. [PubMed: 17336145]
- Fujioka T, Ross B, Okamoto H, Takeshima Y, Kakigi R, Pantev C. Tonotopic representation of missing fundamental complex sounds in the human auditory cortex. *Eur J Neurosci*. 2003; 18:432–40. [PubMed: 12887425]
- Gunji A, Hoshiyama M, Kakigi R. Auditory response following vocalization: a magnetoencephalographic study. *Clin Neurophysiol*. 2001; 112:514–20. [PubMed: 11222973]
- Hain TC, Burnett TA, Kiran S, Larson CR, Singh S, Kenney MK. Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research*. 2000; 130:133–141.
- 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; 40:237–240.
- He C, Trainor LJ. Finding the pitch of the missing fundamental in infants. *J Neurosci*. 2009; 29:7718–8822. [PubMed: 19535583]
- Heinks-Maldonado TH, Mathalon DH, Gray M, Ford JM. Fine-tuning of auditory cortex during speech production. *Psychophysiology*. 2005; 42:180–90. [PubMed: 15787855]

- Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience*. 2002; 14:1125–1138. [PubMed: 12495520]
- Jurgens U. The Neural Control of Vocalization in Mammals: A Review. *Journal of Voice*. 2008
- Kawahara H, Aikawa K. Contributions of auditory feedback frequency components on F0 fluctuations. *Journal of the Acoustical Society of America*. 1996; 100:2825(A).
- Larson CR. On the relation of PAG neurons to laryngeal and respiratory muscles during vocalization in the monkey. *Brain Res*. 1991; 552:77–86. [PubMed: 1913183]
- Liu H, Behroozmand R, Larson CR. Enhanced neural responses to self-triggered voice pitch feedback perturbations. *Neuroreport*. 2010a; 21:527–31. [PubMed: 20386347]
- Liu H, Meshman M, Behroozmand R, Larson CR. Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clinical Neurophysiology*. 2010b
- Ludlow CL. Central nervous system control of the laryngeal muscles in humans. *Respir Physiol Neurobiol*. 2005; 147:205–22. [PubMed: 15927543]
- 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; 24:375–425. [PubMed: 3615753]
- Oostenveld R, Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*. 2001; 112:713–9. [PubMed: 11275545]
- Pantev C, Hoke M, Lutkenhoner B, Lehnertz K. Tonotopic organization of the auditory cortex: pitch versus frequency representation. *Science*. 1989; 246:486–8. [PubMed: 2814476]
- 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; 100:164–70. [PubMed: 8922991]
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The processing of temporal pitch and melody information in auditory cortex. *Neuron*. 2002; 36:767–76. [PubMed: 12441063]
- Penagos H, Melcher JR, Oxenham AJ. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J Neurosci*. 2004; 24:6810–5. [PubMed: 15282286]
- Ragot R, Lepaul-Ercole R. Brain potentials as objective indexes of auditory pitch extraction from harmonics. *Neuroreport*. 1996; 7:905–9. [PubMed: 8724670]
- Ross B, Tremblay K. Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hear Res*. 2009; 248:48–59. [PubMed: 19110047]
- Sivasankar M, Bauer JJ, Babu T, Larson CR. Voice responses to changes in pitch of voice or tone auditory feedback. *Journal of the Acoustical Society of America*. 2005; 117:850–857. [PubMed: 15759705]
- Titze, IR. *Principles of Voice Production*. Prentice hall, Englewood Cliffs; 1994.
- Tomlinson RW, Schwarz DW. Perception of the missing fundamental in nonhuman primates. *J Acoust Soc Am*. 1988; 84:560–5. [PubMed: 3170947]
- Trainor L. Science & music: the neural roots of music. *Nature*. 2008; 453:598–9. [PubMed: 18509429]
- Winkler I, Tervaniemi M, Naatanen R. Two separate codes for missing-fundamental pitch in the human auditory cortex. *Journal of the Acoustical Society of America*. 1997; 102:1072–1082. [PubMed: 9265755]
- Xu Y, Larson C, Bauer J, Hain T. Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences. *Journal of the Acoustical Society of America*. 2004; 116:1168–1178. [PubMed: 15376682]
- Zatorre RJ. Pitch perception of complex tones and human temporal-lobe function. *J Acoust Soc Am*. 1988; 84:566–72. [PubMed: 3170948]
- Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. *Cereb Cortex*. 2001; 11:946–53. [PubMed: 11549617]
- Zatorre RJ. Neuroscience: finding the missing fundamental. *Nature*. 2005; 436:1093–4. [PubMed: 16121160]

- The auditory system can resolve the pitch in sounds with missing fundamental (F0).
- The right auditory cortex is specialized for pitch processing in the absence of F0.
- The speed of pitch processing is faster in the left cortical auditory areas.
- Pitch perturbation in voice auditory feedback triggers compensatory vocal reactions.
- The audio-vocal integration system controls for vocal pitch output during speech.

Figure 1.

a) Time course of the grand averaged ERPs in 14 subjects in response to pitch shifts in complex tone (Tone), missing fundamental complex tone (Tone_{MF}) and pure tone (F0) feedback in the left (C₃), vertex (C_z) and right (C₄) electrodes. **b)** Topographical scalp distributions of the N1 difference maps for Tone-Tone_{MF} in a time window spanning from 95–105 ms post-stimulus.

Figure 2.

a) Time course of the grand averaged ERPs in 14 subjects in response to pitch shifts in voice (Voice), missing fundamental voice (Voice_{MF}) and pure tone (F₀) feedback in the left (C₃), vertex (C_z) and right (C₄) electrodes. **b)** Topographical scalp distributions of the N1 difference maps for Voice-Voice_{MF} in a time window spanning from 95–105 ms post-stimulus.

Figure 3.

Boxplots of **a)** P1, **b)** N1 and **c)** P2 ERP response latencies to pitch shifts in Tone, Tone_{MF} and F0 feedback (left subplot) and Voice, Voice_{MF} and F0 feedback (right subplot) in the left (black boxes) and right (gray boxes) hemispheres. The boxplots represent the median along with 25th and 75th percentiles values and the whiskers extend to the most extreme data point not considered an outlier, and the outliers are plotted as individual points. The “+” signs in each boxplot mark the mean of the ERP response latencies.

Figure 4.

Boxplots of **a)** P1, **b)** N1 and **c)** P2 ERP response amplitudes to pitch shifts in Tone, Tone_{MF} and F0 feedback (left subplot) and Voice, Voice_{MF} and F0 feedback (right subplot) in the left (black boxes) and right (gray boxes) hemispheres. The boxplots represent the median along with 25th and 75th percentiles values and the whiskers extend to the most extreme data point not considered an outlier, and the outliers are plotted as individual points. The “+” signs in each boxplot mark the mean of the ERP response amplitudes.

Figure 5.

Time course of the grand averaged vocal responses in 14 subjects in response to pitch shifts in voice (Voice), missing fundamental voice (Voice_{MF}), complex tone (Tone), missing fundamental complex tone (Tone_{MF}) and pure tone (F0) feedback. The horizontal and vertical dotted lines mark the baseline and pitch shift stimulus onset, respectively.

Figure 6.

Schematic of the experimental setup. ERPs were obtained in response to +200 cents pitch shift stimulus (PSS) in a randomly chosen auditory feedback including subject's own voice (Voice), missing fundamental voice (Voice_{MF}: Voice-F₀), periodic complex tone (Tone: F₀+H₁+H₂+H₃+H₄), missing fundamental complex tone (Tone_{MF}: H₁+H₂+H₃+H₄) and pure tone (F₀). In order to generate complex tones, subject's voice signal was fed to a Kay Visipitch for online pitch extraction and the analog F₀ output was used to drive five oscillators that generated sinusoidal tones at voice F₀ and its first (H₁), second (H₂), third (H₃) and fourth (H₄) harmonic frequencies. To create missing fundamental feedback, voice and complex tone channels were high-pass filtered (HPF) at 1.5×F₀ frequency. The voice signal was also fed to a voice onset detector (VOD) module that was used to cue the pitch shifter to deliver PSS after a randomly chosen time delay (Δt) between 500–1000 ms following vocal onset. The figure illustrates an example of a condition where a pure sinusoidal tone at F₀ is randomly chosen as voice feedback during vocalization of the vowel sound /a/.

Figure 7.

The frequency spectra and time waveforms of voice (Voice), missing fundamental voice (Voice_{MF}), complex tone (Tone), missing fundamental complex tone (Tone_{MF}) and pure tone (F0) feedback during 100 ms vocalization of the vowel sound /a/ in a subject with a conversational F0 at about 195 Hz.