

Published in final edited form as:

Clin Neurophysiol. 2011 May ; 122(5): 951–957. doi:10.1016/j.clinph.2010.08.010.

Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback

Hanjun Liu^a, Michelle Meshman^b, Roozbeh Behroozmand^b, and Charles R. Larson^b

^aDepartment of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, 510080, P. R. China

^bDepartment of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL, 60208

Abstract

Objective—The present study examined the differential effects of voice auditory feedback perturbation direction and magnitude on voice fundamental frequency (F_0) responses and event-related potentials (ERPs) from EEG electrodes on the scalp.

Methods—The voice F_0 responses and N1 and P2 components of ERPs were examined from twelve right-handed speakers when they sustained a vowel phonation and their mid-utterance voice pitch feedback was shifted ± 100 , ± 200 , and ± 500 cents with 200 ms duration.

Results—Downward voice pitch feedback perturbations led to larger voice F_0 responses than upward perturbations. The amplitudes of N1 and P2 components were larger for downward compared with upward pitch shifts for 200 and 500 cents stimulus magnitudes. Shorter N1 and P2 latencies were also associated with larger magnitudes of pitch feedback perturbations.

Conclusion—Corresponding changes in vocal and neural responses to upward and downward voice pitch feedback perturbations suggest that the N1 and P2 components of ERPs reflect neural concomitants of the vocal responses.

Significance—The findings of interactive effects between the magnitude and direction of voice feedback pitch perturbation on N1 and P2 ERP components indicate that the neural mechanisms underlying error detection and correction in voice pitch auditory feedback are differentially sensitive to both the magnitude and direction of pitch perturbations.

Keywords

Auditory feedback; N1-P2 complex; Event-related potential (ERP); Pitch feedback perturbation; Vocalization

1. Introduction

In recent years, the frequency perturbation paradigm has been developed to investigate the role of auditory feedback for the control of voice fundamental frequency (F_0) (Elman, 1981,

© 2010 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

Correspondence to Dr. Charles R. Larson, PhD, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, 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.

Kawahara, 1994, Burnett, et al., 1998, Larson, 1998, Jones and Munhall, 2002). In this paradigm, speakers vocalize a vowel sound or a speech phrase, during which their voice pitch feedback is unexpectedly shifted and fed back to them through the headphones. In response, speakers change their voice F_0 in the direction opposite to the perturbation. That is, they lower their voice F_0 when the pitch feedback is shifted upwards, and vice versa. This compensatory response to pitch feedback perturbations has been observed not only during vowel phonation tasks (Burnett, et al., 1998, Hain, et al., 2000, Larson, et al., 2000, Liu and Larson, 2007, Larson, et al., 2008) but also during speech production (Natke and Kalveram, 2001, Donath, et al., 2002, Natke, et al., 2003, Xu, et al., 2004, Chen, et al., 2007, Liu, et al., 2009), suggesting that auditory feedback can be used to stabilize the voice F_0 around the desired level and facilitate the accurate and timely control of speech production.

Recently, event-related potential (ERP) and magnetoencephalography (MEG) techniques along with the frequency perturbation paradigm have been used to explore the temporal neural processing of auditory feedback (Heinks-Maldonado, et al., 2005, Heinks-Maldonado, et al., 2006, Behroozmand, et al., 2009, Hawco, et al., 2009, Behroozmand, et al., in press). In these studies, the N1/M1/P2 responses were primarily used to reflect neural changes induced by active vocalization in comparison with passive listening. It was found that N1/M1 responses during active vocalization were significantly suppressed relative to those during passive listening when pitch feedback perturbations were presented at vocal onset (Heinks-Maldonado, et al., 2005, Heinks-Maldonado, et al., 2006), consistent with results reported by Houde et al. (2002). In contrast, when pitch-shifted stimuli were delivered with delays relative to the vocal onset, P2 responses during active vocalization were enhanced as compared to those during passive listening (Behroozmand, et al., 2009, Behroozmand, et al., in press).

By comparing the difference in the neural processing of speaking and listening, these pitch-shift studies demonstrate that auditory feedback plays a crucial role in differentiating self-produced voice from an alien voice or external sound. How the audio-vocal system uses auditory feedback to regulate the voice (e.g. increase or decrease voice F_0) to reach the communication target during vocalization, however, is still unclear. Two recent ERP studies showed that greater neural responses were associated with larger pitch-shift stimuli when subjects were asked to sustain a vowel phonation (Behroozmand, et al., 2009, Hawco, et al., 2009). It is noteworthy that, in these two studies and other related studies (Heinks-Maldonado, et al., 2005, Heinks-Maldonado, et al., 2006), only upward pitch-shift stimuli were presented during testing. However, it is well known that people compensate for voice feedback errors by changing their voice F_0 in the direction opposite to the pitch-shift stimulus regardless of stimulus direction (Burnett, et al., 1998, Larson, 1998), indicating that the audio-vocal system can detect the direction of pitch feedback perturbations and adjust its responses accordingly. Although the response direction depends on the stimulus direction, vocal response magnitudes and latencies do not differ with stimulus direction during sustained vowels (Larson, et al., 2001, Chen, et al., 2007). It is unknown if ERP responses vary with stimulus direction since this variable has not previously been tested. In addition, ERP amplitudes have been shown to differ when the magnitude of the pitch-shift stimulus was varied (Behroozmand, et al., 2009, Hawco, et al., 2009), whereas vocal response magnitudes were not systematically modulated as a function of these magnitudes (Burnett, et al., 1998, Chen, et al., 2007). This lack of a correspondence between ERP and vocal response magnitudes to stimulus magnitude and direction in previous studies may be related to different procedures in the behavioral and the electrophysiological studies.

Therefore, the present study was designed to investigate vocal and neural responses to voice pitch-shifted feedback using the same procedures that were previously used only in the behavioral studies. The properties of pitch-shift stimuli were manipulated by stimulus

magnitude (100, 200, and 500 cents) and stimulus direction (upward and downward) while the duration was set to 200 ms. By examining the vocal as well as the N1 and P2 responses to voice feedback perturbations, the present study was designed to determine if there are corresponding changes in these responses to variations in the stimuli. Determining such correspondences would help to further our understanding of the neural mechanisms related to the role of auditory feedback in vocal control.

2. Method and Materials

2.1 Subjects

Native English-speaking, right-handed students (3 male and 9 female; ages 19-22) from Northwestern University participated in the experiment. None of the subjects reported a history of speech, hearing, or neurological disorders. All signed informed consent approved by the Northwestern University Institutional Review Board and were paid for their participation.

2.2 Apparatus

During the experiment, subjects wore an AKG microphone (model C420) in a sound booth. Etymotic insert earphones (model ER1-14A) were inserted bilaterally into the subjects' ear canals. Subjects were asked to vocalize the vowel /a/ at approximately 70 dB SPL and at their normal conversational pitch. The voice signal from the microphone was first amplified with a Mackie mixer (model 1202) and shifted in pitch with an Eventide Eclipse Harmonizer. The voice signal was further amplified with a Mackie mixer (model 1202-VZL) and finally amplified with a Crown D75 amplifier and HP 350 dB attenuators at 80 dB SPL. MIDI software (Max/MSP v.4.6 by Cycling 74) was used to control the harmonizer. The MIDI software also generated a TTL pulse to mark the onset of each stimulus for synchronized averaging of the recorded ERPs. Prior to the experiments, a Zwislocki coupler and Brüel & Kjær sound level meter (model 2250) were used for calibration. These calibration procedures were done to make sure there was a gain of 10 dB SPL between the subject's voice amplitude and feedback loudness.

Subjects wore an EEG cap (Electro-Cap International, Inc.) consisting of Ag-AgCl electrodes, in which the electrodes are distributed according to the international 10-20 system of electrode placement. Auditory ERPs were collected from 7 sites distributed bilaterally on the subjects' scalp (Cz, C4, C3, F4, F3, F8, F7). These electrodes were referenced to earlobe-linked electrodes with a ground electrode on the forehead. An electrode impedance meter (Grass, EZM 5AB) was used to ensure that the impedance between the electrode and reference was lower than 5k Ohms at 30 Hz. Auditory EEGs were amplified with gain of 10k (Grass P511 AC amplifier). Voice, feedback, TTL pulses, and EEG signals were digitized by PowerLab (10 kHz, 12 bit, AD Instruments) and recorded on a laboratory computer utilizing Chart software (AD Instruments).

2.3 Procedures

The experiment consisted of six blocks of trials: two stimulus directions (upward and downward) by three stimulus magnitudes (100, 200, and 500 cents). During each trial, subjects were instructed to vocalize the vowel sound /a/ for about 6-7 seconds while their voice pitch was shifted 6 times and fed back to them through the headphones. During each vocalization, the direction and magnitude of the pitch-shift stimuli were held constant. Between each vocalization, a short break of 3-5 seconds was taken by the subjects to avoid vocal fatigue. The duration of each pitch-shift stimulus was fixed at 200 ms, and the inter-stimulus interval varied between 700-900 ms. In a single block, subjects were asked to produce twenty-five consecutive vocalizations, yielding a total of 150 trials (25

vocalizations \times 6 trials per vocalization) of the same stimulus magnitude and direction. Subjects were asked to start their first vocalization when signaled by hand from the experimenter. For the rest of their vocalizations, they were told to take a break and then start vocalizing when they felt ready. The sequence of the stimulus properties (magnitude and direction) was randomized across blocks of trials and subjects.

2.4 Behavioral Data Analysis

Event-related averaging techniques were used to measure the magnitude and latency of vocal responses (Chen, et al., 2007, Larson, et al., 2008). First, the voice signal was processed in Praat (Boersma, 2001) to produce a train of pulses corresponding to the fundamental period of the voice waveform. This pulse train was transformed to an F_0 contour wave in IGOR PRO (v.6.0 by Wavemetrics, Inc., Lake Oswego, OR), and then converted to a cent scale using the following:

$$\text{Voice(cent)}=100 \times (12 \times \log(f2/f1)/\log(2))$$

where $f1$ is an arbitrary reference note at 195.997 Hz (G4), and $f2$ is the voice F_0 in Hertz.

The voice contours were then time-aligned with the TTL pulse with a window of 200 ms pre- and 500 ms post-stimulus. Prior to the averaging, visual inspections were performed on each individual trial so that those trials with very large F_0 fluctuations resulting from either signal processing errors or vocal interruption were removed from further analysis. In addition, those vocal trials associated with bad EEG trials were also rejected (see below). Averaged F_0 contours that exceeded a value of two standard deviations (SDs) of the pre-stimulus mean (baseline F_0) beginning at least 60 ms after the stimulus and lasting at least 50 ms were considered to be valid. Response latency was measured as the time from the stimulus onset at which the response exceeded 2 SDs of the pre-stimulus mean, and response magnitude was measured as the difference between the pre-stimulus mean and the largest value of the F_0 contour following the response onset. A response was considered as “opposing” if its direction was opposite to the stimulus direction and “following” if it was in the same direction as the stimulus. Vocal response magnitude and latency were measured and analyzed only for the opposing responses in each condition.

2.5 EEG Data Analysis

Off-line filtering was performed on the EEG signals using a band-pass filter with cut-off frequencies set to 1 and 30 Hz. The filtered EEG signal was then cut into epochs ranging from -200 ms to 700 ms, relative to the onset of the stimulus. Analysis procedures performed in Igor PRO rejected trials in which ERP artifacts exceeded $\pm 50 \mu\text{V}$ prior to averaging; the voice F_0 signals corresponding to these trials were also rejected. On average, 16% of the trials across all subjects were rejected. The pre-stimulus mean amplitude of the remaining epochs was calculated and subtracted from the signals prior to averaging. The N1 and P2 ERP components, including the amplitudes and latencies, were extracted from the averaged neural responses by finding the most prominent peaks in 50 ms-long time windows centered at 100 ms and 200 ms, respectively. Values of amplitudes and latencies of N1 and P2 responses were put into SPSS (v. 16.0) to test for significant differences across the conditions using repeated-measures analysis of variance (RMANOVA). Appropriate t-tests were calculated when significant interactions were observed. Probability values were corrected for multiple degrees of freedom using Greenhouse-Geisser if the assumption of sphericity was violated. Corrected p values were reported along with original degrees of freedom.

3. Results

Behavioral Results

Seventy-two averaged responses (12 subjects \times 3 stimulus magnitudes \times 2 stimulus directions), including “following”, opposing, and non-responses, were obtained (see Table 1). It can be seen that 13 of 72 responses “followed” the direction of stimulus, and the majority of the “following” responses (8 of 13) occurred under the 500 cents condition. Four of 72 responses did not meet the criteria of valid responses and were declared to be non-responses. Figure 1 shows the vocal responses averaged across all subjects as a function of stimulus magnitude and direction (down stimuli at top, and up stimuli at bottom), where it can be seen that larger response magnitudes were associated with downward pitch-shift stimuli as compared to upward stimuli.

Although these data lend themselves to repeated-measures ANOVAs (Xu, et al., 2004), two-way factorial ANOVAs were used because the assumptions of repeated-measures ANOVA were not met due to missing data and unequal cell size. For the vocal response magnitude, a significant main effect was found for stimulus direction ($F(1, 49)=6.084, p=0.017$) but not for stimulus magnitude ($F(2, 49)=0.270, p=0.765$) or stimulus magnitude \times direction interaction ($F(2, 49)=0.293, p=0.747$). Downward pitch-shift stimuli elicited significantly larger responses (9.2 ± 5.1 cents) than upward stimuli (6.4 ± 5.2 cents). With regard to the latencies of vocal responses, statistical results showed no main effects of stimulus magnitude ($F(2, 49)=0.156, p=0.856$) or stimulus direction ($F(1, 49)=3.313, p=0.075$). No significant interaction between these two variables was found ($F(2, 49)=0.032, p=0.968$).

EEG Results

Figures 2-3 show ERP waveforms averaged over all subjects as a function of stimulus magnitude and stimulus direction. It can be seen that a negative peak and a positive peak are located at around 100 ms and 200 ms, which are termed N1 and P2 in the present study. As shown in Figure 2, the response latencies for 500 cents pitch-shifted stimuli are shorter than those for 100 and 200 cents stimuli across the electrodes. Figure 4 shows the N1 (top) and P2 (middle) amplitudes and latencies (bottom) across stimulus conditions. Since no main effects of stimulus direction were found on the N1 and P2 latencies (see below), values of latencies to these two directions were collapsed and are not shown in Figure 4.

Three-way (3 stimulus magnitudes, 2 stimulus directions, and 7 electrode positions) RMANOVAs were performed on the N1 and P2 amplitudes, respectively. For the N1 amplitudes, statistical results showed significant main effects of electrode position ($F(6, 66)=7.088, p<0.001$), stimulus magnitude ($F(2, 22)=2.015, p=0.004$), and stimulus direction ($F(1, 11)=4.963, p=0.023$). A significant interaction was observed between stimulus direction and stimulus magnitude ($F(2, 22)=2.055, p=0.006$). Further t-tests showed significantly larger responses for 200 ($t=-3.314, df=166, p=0.001$) and 500 cents ($t=-4.166, df=166, p<0.001$) than those for 100 cents in the case of downward pitch-shift stimuli. No significant difference was found between 200 and 500 cents ($t=-1.179, df=166, p=0.240$). This effect of stimulus magnitude on the N1 amplitudes, however, was not observed for upward stimuli. In addition, larger N1 amplitudes to downward stimuli relative to upward stimuli were observed for 200 cents ($t=2.330, df=166, p=0.021$) and 500 cents ($t=-2.538, df=166, p=0.012$), but not for the 100 cents condition ($t=1.532, df=166, p=0.127$).

For the P2 amplitudes, significant main effects were observed for electrode position ($F(6, 66)=14.966, p<0.001$), stimulus magnitude ($F(2, 22)=3.258, p=0.001$), and stimulus direction ($F(1, 11)=8.127, p=0.002$), as well as a significant stimulus direction \times stimulus magnitude ($F(2, 22)=3.947, p=0.016$) interaction. Further t-tests revealed, for the downward direction, significant differences in the P2 amplitudes between 200 and 500 cents ($t=-2.474,$

df=166, $p=0.014$), between 100 and 500 cents ($t=-3.157$, $df=166$, $p=0.002$), but not between 100 and 200 cents ($t=-0.946$, $df=166$, $p=0.346$). The greatest P2 amplitudes were associated with 500 cents, followed by 200 cents and 100 cents. In contrast, stimulus magnitude had no effect on the P2 amplitudes for the upward stimuli. In addition, larger P2 amplitudes to downward stimuli relative to upward stimuli were observed for 500 cents ($t=3.275$, $df=166$, $p=0.001$) but not for the 100 cents ($t=0.397$, $df=166$, $p=0.692$) and 200 cents ($t=1.325$, $df=166$, $p=0.187$) conditions.

Three-way RMANOVAs on N1 latencies showed significant main effects for stimulus magnitude ($F(2, 22)=41.005$, $p<0.001$) and electrode position ($F(6, 66)=4.469$, $p=0.001$) but not for stimulus direction ($F(1, 11)=0.772$, $p=0.398$). No significant interactions among these variables were found. Post-hoc Bonferroni tests revealed that the longest N1 latencies were associated with 100 cents (145 ms), followed by 200 cents (134 ms) and 500 cents stimuli (123 ms), in which each of the three stimuli was significantly different from other two stimuli ($p<0.002$).

Three-way RMANOVAs on P2 latencies showed significant main effects of stimulus magnitude ($F(2, 22)=22.297$, $p<0.001$) and electrode position ($F(6, 66)=5.617$, $p<0.001$), but no main effect of stimulus direction ($F(1, 11)=0.599$, $p=0.455$) or significant interactions among these three variables were found. Post-hoc Bonferroni tests revealed that 100 cents induced the longest latencies (236 ms) compared to 200 cents (217 ms, $p=0.007$) and 500 cents stimuli (210 ms, $p<0.001$), and no significant differences were found in the P2 latency between 200 cents and 500 cents stimuli ($p=0.296$).

4. Discussion

The present study was designed to address the neural processing of pitch perturbation in voice auditory feedback as a function of perturbation direction and magnitude. Results showed that downward pitch-shift stimuli elicited larger N1/P2 amplitudes than upward stimuli for 200 and 500 cents stimulus magnitudes. Vocal responses were also larger for downward compared to upward stimuli. These findings demonstrate that there are differential effects of perturbation magnitude and direction on the neural processing of voice pitch feedback during vocalization.

Our findings indicated that larger pitch feedback perturbation magnitudes elicited larger N1/P2 amplitudes, which is complementary to the findings of larger neural responses associated with larger pitch feedback perturbations in recent pitch-shift ERP studies (Behroozmand, et al., 2009, Hawco, et al., 2009). However, in the present study the effect of pitch perturbation magnitude on the ERPs was only observed when pitch feedback was shifted downward, while in the two previous studies this effect occurred with upward pitch-shifted stimuli. Although this effect may just be due to sampling error, it is worthwhile considering other explanations. For example, the significant procedural differences between the Hawco et al. (2009) and the present study, such as different stimulus magnitudes (0, 25, 50, 100, and 200 cents), the requirement of matching a music note during vocalization, the use of babble masking noise, and the measurement of ERPs using the MMN, are all factors that could have led to these differences. It is more difficult to explain the effect of stimulus direction by comparison with the Behroozmand et al. (2009) study, since both studies used a similar paradigm except the stimulus direction. It could be partially related to the task difference between these two studies: the Behroozmand et al. (2009) study consisted of vocalization and subsequent listening tasks, while only the vocalization task was involved in the present study.

In this discussion it is also important to note that the downward stimulus in comparison to the upward direction led to larger ERP responses as well as voice F_0 responses. Considering methodological differences between the present and previous behavioral studies may provide additional insights into the causes of the directional effect observed in the present study. Although several previous pitch-shift studies manipulated the direction of pitch feedback perturbation (Larson, et al., 2001, Chen, et al., 2007, Liu and Larson, 2007, Larson, et al., 2008), the directional effect on vocal response magnitudes has rarely been reported. In each block of trials in the present study, voice pitch feedback was shifted with a fixed direction (either upward or downward) 150 times, while in previous behavioral studies, upward or downward pitch-shift stimuli were randomly intermixed a total of 20-60 times. Presenting 150 trials in which the pitch-shift direction is always the same may generate expectations about the sequence of stimuli that causes downward stimuli to elicit larger behavioral and neural responses than upward stimuli. Whether different presentation sequences of pitch feedback perturbations (i.e. shifted with a random or fixed sequence of stimulus direction) influence the voice F_0 regulation is also unclear. Future behavioral or neural pitch-shifted studies should be conducted to explore these questions.

If future studies can verify that stimulus direction does indeed lead to differences in ERP responses, they would provide direct evidence for different neural processing mechanisms for upward vs. downward pitch-shifted voice feedback. We know that such mechanisms must exist in order for the audio-vocal system to be able to detect differences in stimulus direction and generate vocal responses in the appropriate direction. Further studies should be conducted to explore the specific neural mechanisms in the processing of upward and downward pitch feedback perturbations during vocalization.

It is noteworthy that stimulus magnitude and direction did not affect the neural processing of pitch feedback perturbation independently. That is, a significant interaction was observed between stimulus magnitude and stimulus direction, where downward pitch-shifted stimuli elicited larger N1/P2 amplitudes in the case of the 500 cents condition, but not for the 100 cents condition. Also, the modulation of N1/P2 amplitudes as a function of stimulus magnitude was found for the downward stimuli but not for upward stimuli. All these findings indicate that different stimulus properties (e.g. magnitude, direction) interact with each other on the ERP amplitudes and latencies, suggesting that the effects of these variables on neural processing of pitch feedback perturbations during vocalization may be more complex than previously thought.

5. Conclusion

The present study examined the N1 and P2 components of ERPs related to voice pitch feedback perturbations during vocalization. The results showed that, when voice pitch feedback was shifted downward, N1/P2 amplitudes increased along with the increase in stimulus magnitude, and the responses were larger than for upward stimuli. Behaviorally larger vocal response magnitudes were also more frequently associated with downward than with upward stimuli. In addition, N1 and P2 latencies decreased with the increase in the magnitude of pitch feedback perturbations. These findings suggest that the N1-P2 complex is sensitive to the direction of voice pitch-shifted feedback and may reflect different mechanisms for detecting and responding to upwards vs. downwards changes in voice pitch feedback.

Acknowledgments

This work was supported by NIH Grant No.1R01DC006243 and National Science Foundation of China Grant No. 30970965. The authors thank Chun Liang Chan for programming assistance.

References

- Behroozmand R, Karvelis L, Liu H, Larson CR. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clin Neurophysiol.* 2009; 120:1303–1312. [PubMed: 19520602]
- Behroozmand R, Liu H, Larson CR. Time-dependent neural processing of the auditory feedback during voice pitch error perturbation. *J Cogn Neurosci.* in press.
- Boersma P. Praat, a system for doing phonetics by computer. *Glott International.* 2001; 5(9/10):341–345.
- Burnett TA, Freedland MB, Larson CR, Hain TC. Voice F0 Responses to Manipulations in Pitch Feedback. *J Acoust Soc Am.* 1998; 103:3153–3161. [PubMed: 9637026]
- Chen SH, Liu H, Xu Y, Larson CR. Voice F0 responses to pitch-shifted voice feedback during English speech. *J Acoust Soc Am.* 2007; 121:1157–1163. [PubMed: 17348536]
- Donath TM, Natke U, Kalveram KT. Effects of frequency-shifted auditory feedback on voice F0 contours in syllables. *J Acoust Soc Am.* 2002; 111:357–366. [PubMed: 11831808]
- Elman JL. Effects of frequency-shifted feedback on the pitch of vocal productions. *J Acoust Soc Am.* 1981; 70:45–50. [PubMed: 7264071]
- 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. *Exp Brain Res.* 2000; 130:133–141. [PubMed: 10672466]
- Hawco CS, Jones JA, Ferretti TR, Keough D. ERP correlates of online monitoring of auditory feedback during vocalization. *Psychophysiology.* 2009; 46:1216–1215. [PubMed: 19674393]
- Heinks-Maldonado TH, Mathalon DH, Gray M, Ford JM. Fine-tuning of auditory cortex during speech production. *Psychophysiology.* 2005; 42:180–190. [PubMed: 15787855]
- Heinks-Maldonado TH, Nagarajan SS, Houde JF. Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport.* 2006; 17:1375–1379. [PubMed: 16932142]
- Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: An MEG study. *J Cogn Neurosci.* 2002; 14:1125–1138. [PubMed: 12495520]
- Jones JA, Munhall KG. The role of auditory feedback during phonation: studies of Mandarin tone production. *J Phon.* 2002; 30:303–320.
- Kawahara H. Interactions between speech production and perception under auditory feedback perturbations on fundamental frequencies. *J Acoust Soc Jpn.* 1994; 15:201–202.
- Larson CR. Cross-modality influences in speech motor control: The use of pitch shifting for the study of F0 control. *J Commun Disord.* 1998; 31:489–503.
- Larson CR, Altman KW, Liu H, Hain TC. Interactions between auditory and somatosensory feedback for voice F (0) control. *Exp Brain Res.* 2008; 187:613–621. [PubMed: 18340440]
- Larson CR, Burnett TA, Bauer JJ, Kiran S, Hain TC. Comparisons of voice F0 responses to pitch-shift onset and offset conditions. *J Acoust Soc Am.* 2001; 110:2845–2848. [PubMed: 11785786]
- Larson CR, Burnett TA, Kiran S, Hain TC. Effects of pitch-shift onset velocity on voice F0 responses. *J Acoust Soc Am.* 2000; 107:559–564. [PubMed: 10641664]
- Liu H, Larson CR. Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *J Acoust Soc Am.* 2007; 122:3671–3677. [PubMed: 18247774]
- Liu H, Xu Y, Larson CR. Attenuation of vocal responses to pitch perturbations during Mandarin speech. *J Acoust Soc Am.* 2009; 125:2299–2306. [PubMed: 19354405]
- Natke U, Donath TM, Kalveram KT. Control of voice fundamental frequency in speaking versus singing. *J Acoust Soc Am.* 2003; 113:1587–1593. [PubMed: 12656393]
- Natke U, Kalveram KT. Effects of frequency-shifted auditory feedback on fundamental frequency of long stressed and unstressed syllables. *J Speech Lang Hear Res.* 2001; 44:577–584. [PubMed: 11407562]
- Xu Y, Larson C, Bauer J, Hain T. Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences. *J Acoust Soc Am.* 2004; 116:1168–1178. [PubMed: 15376682]

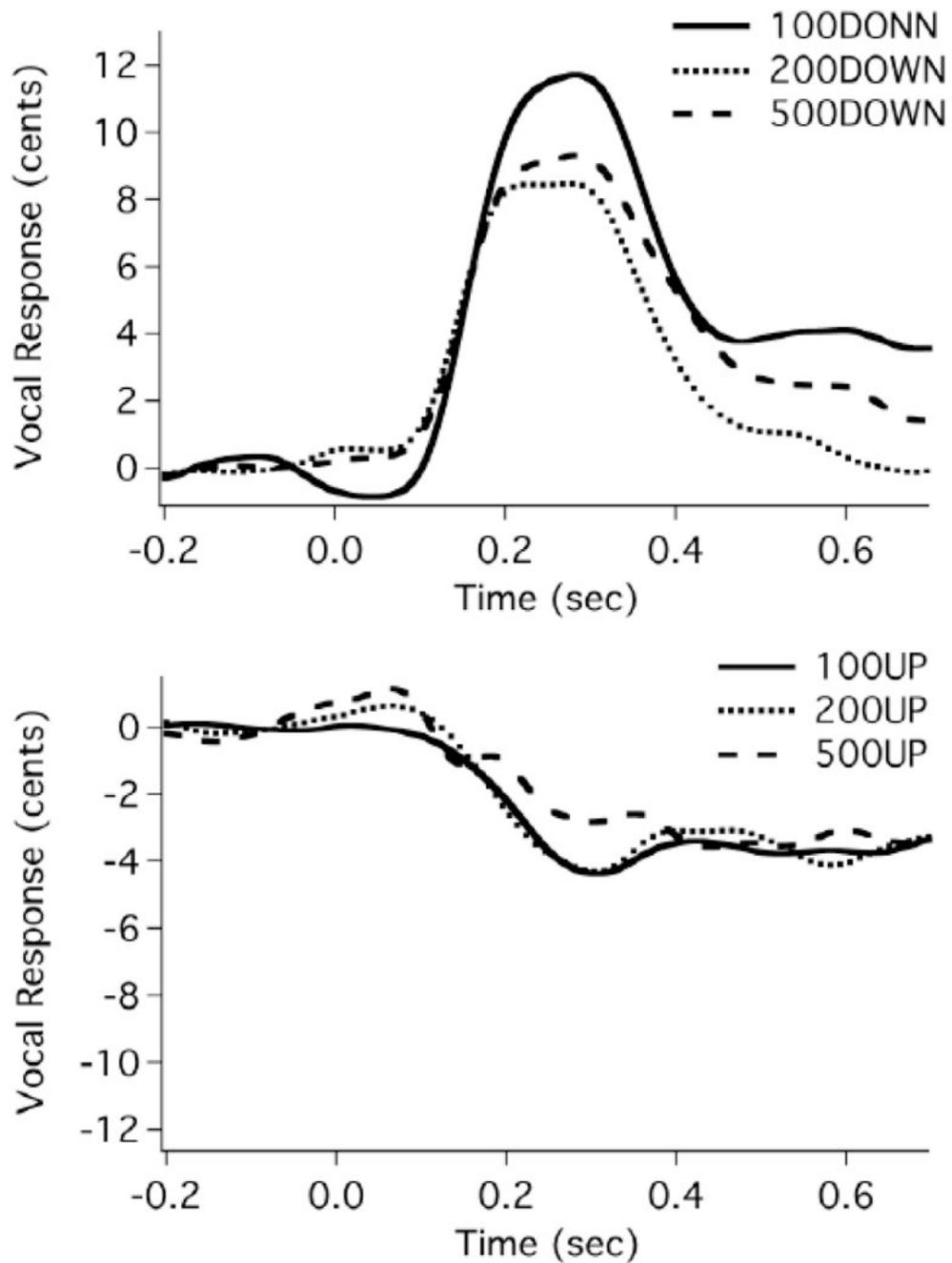


Figure 1.

Composite averages of all compensatory vocal responses to downward (top) and upward stimuli (bottom) for all subjects across stimulus magnitudes. The solid lines, the thick dashed lines, and the sparse dashed lines represent the vocal responses to 100, 200, and 500 cents pitch-shift stimuli, respectively. The stimulus was presented at time 0. Horizontal axis is time in seconds, and vertical axis is voice frequency in cents.

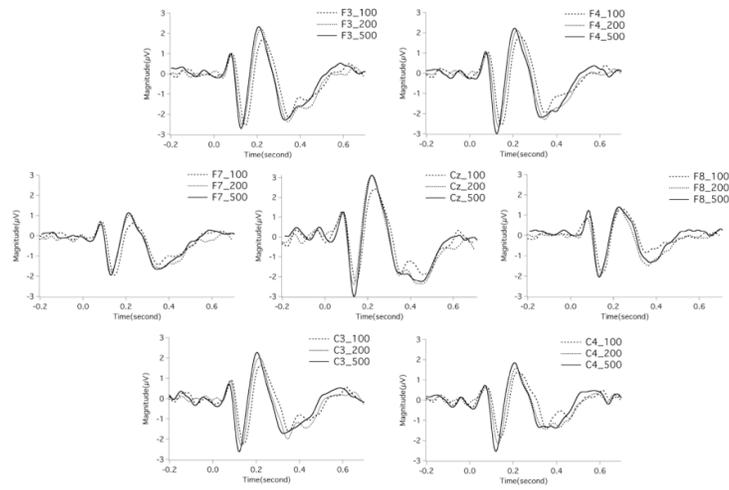


Figure 2. Grand-average event-related potentials (ERPs) as a function of stimulus magnitude at electrodes Cz, C4, C3, F4, F3, F8, and F7, in which the ERPs to upward and downward stimuli were collapsed. The sparse dashed line, the thick dashed line, and the solid line represent the ERPs to 100, 200, and 500 cents pitch perturbations, respectively. The stimulus was presented at time 0.

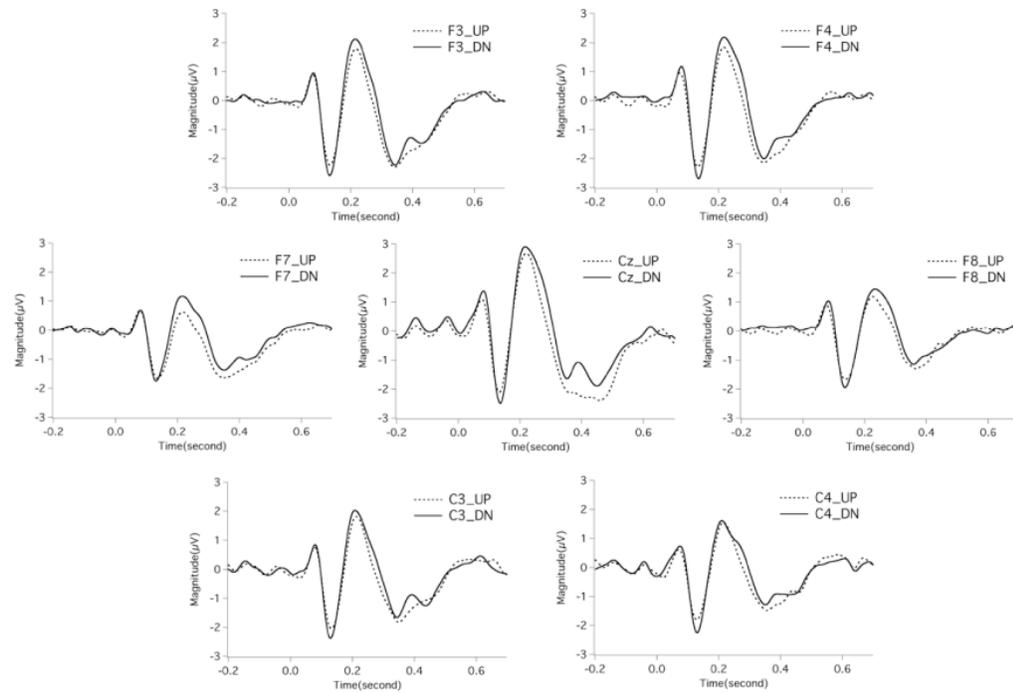


Figure 3.

Grand-average event-related potentials (ERPs) as a function of stimulus direction at electrodes Cz, C4, C3, F4, F3, F8, and F7, in which the ERPs were averaged across stimulus magnitude. The dashed line and the solid line represent the ERPs to upward and downward pitch perturbations, respectively. The stimulus was presented at time 0.

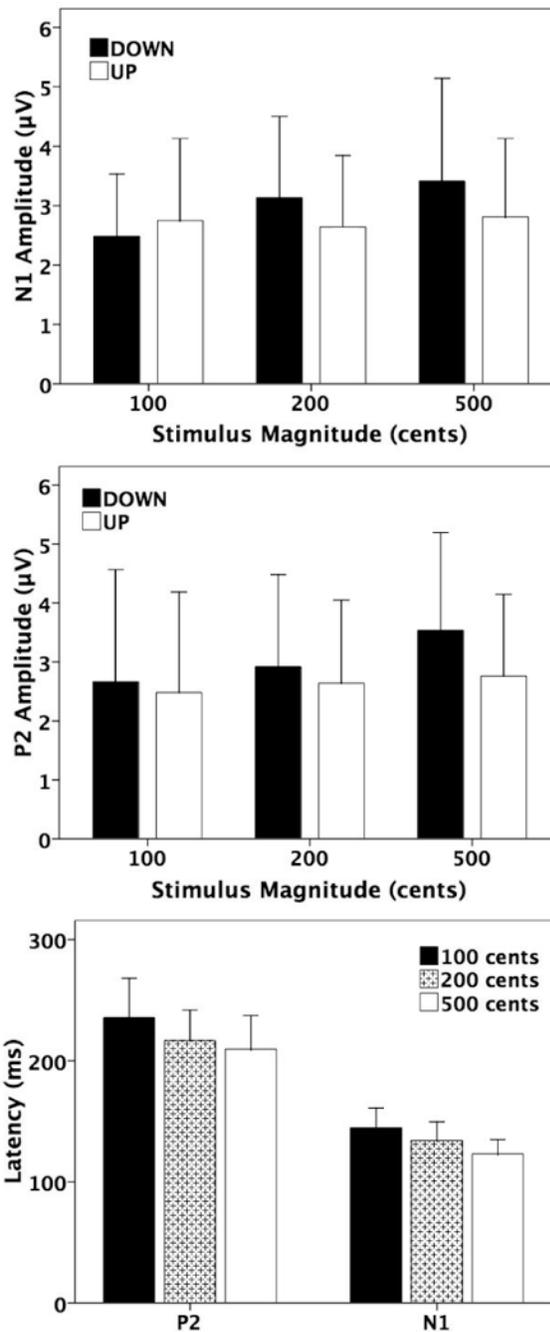


Figure 4.

Effects of magnitude and direction of pitch-shift stimuli on the N1-P2 complex. The T-bar plots on the top and in the middle show the N1 and P2 amplitudes as a function of stimulus magnitude and direction. The black and the white T-bars denote the neural responses to downward and upward pitch-shift stimuli. On the bottom, the T-bar plots represent the N1 and P2 latencies as a function of stimulus magnitude, where the black, the sparse dotted, and the white T-bars denote the neural response to 100, 200, and 500 cents pitch-shift stimuli.

Table 1

The number of opposing (OPP), following (FOL), and non-response (NR) as a function of stimulus direction and stimulus magnitude.

	OPP	FOL	NR	Total
100 cents	21	2	1	24
200 cents	21	3	0	24
500 cents	13	8	3	24
UP	27	8	1	36
DOWN	28	5	3	36