Enhanced neural responses to self-triggered voice pitch feedback perturbations

Hanjun Liu[a], **Roozbeh Behroozmand**[b], and **Charles R. Larson**[b]

[a]Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, PR China

[b]Department of Communication Science and Disorders, Northwestern University, Evanston, Illinois, USA

Abstract

This study investigated the effect of self-triggered voice fundamental frequency (F_0) feedback perturbation on auditory event-related potentials (ERPs) during vocalization and listening. Auditory ERPs were examined in response to self-triggered and computer-triggered −200 cents pitch-shift stimuli while participants vocalized or listened to the playback of their self-vocalizations. The stimuli were either presented with a delay of 500–1000 ms after the participants pressed a button or delivered by a computer with an interstimulus interval of 500–1000 ms. Results showed that self-triggered stimuli elicited larger ERPs compared with computer-triggered stimuli during both vocalization and listening conditions. These findings suggest that self-triggered perturbation of self-vocalization auditory feedback may enhance auditory responses to voice feedback pitch perturbation during vocalization and listening.

Keywords

auditory feedback; event-related potential; fundamental frequency; pitch feedback perturbation; self-triggered stimulation; voice control

Introduction

It is important for the human brain to discriminate self-produced stimuli from externally generated stimuli to avoid the misattribution of the agent of an act. It has been shown that self-triggered motor activity is processed differently from externally triggered stimulation. For example, the phenomenon of motor-induced suppression (MIS) has been observed in the somatosensory system, where brain responses to self-triggered tactile stimulation are suppressed relative to externally triggered tactile stimuli [1,2]. Similar suppression effects have also been found in the auditory cortex. Schafer and Marcus [3] reported suppressed vertex (Cz) electroencephalogram (EEG) responses to self-triggered auditory stimuli relative to externally triggered stimuli. Martikainen *et al.* [4] also reported similar results in one magnetoencephalography study, where the magnetoencephalography responses to self-triggered tones in the auditory cortex were attenuated. In these studies, auditory stimuli were presented immediately after the button press, and therefore the brain had exact foreknowledge of stimulus onset time. Thus, according to the internal forward model theory

Correspondence to Dr Charles R. Larson, PhD, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA, Tel: + 1 847 491 2424; fax: + 1 847 491 4975; clarson@northwestern.edu.

[5], a precise prediction of the actual feedback was available, generating a small prediction error between the intended feedback and the actual feedback, which in turn led to the MIS of cortical responses in the auditory cortex.

Recently, two studies have been performed to investigate the modulation of MIS in the auditory cortex when the timing of stimulus presentation was variable [6,7]. Bass *et al.* [7] compared the auditory event-related potentials (ERPs) in response to self-triggered tones with those elicited by identical tones triggered externally, during which frequency and/or onset of the tone were varied. Results showed that although MIS was observed across all conditions, smaller MIS effects were found for the predictable tone onsets than for the unpredictable onsets. Aliu *et al.* [6] did a similar study, in which tones were presented to the participants either immediately or with a certain delay after the stimulus was self-triggered using a button press. Results indicated that MIS at zero delay was observed only in the left hemisphere, whereas such suppression at non-zero delay was developed bilaterally. In addition, it has been found that auditory cortical responses to self-produced speech are suppressed as compared with the playback of tape-recorded speech [8-11], indicating that speaking, as a specific self-triggered motor act, can also lead to the suppression of brain responses in the auditory cortex relative to listening.

It is noteworthy that the MIS effect during vocal production [8-10] has been reported for conditions in which the stimulus presentation occurred at the onset of vocalization. However, it was found that when the stimulus was presented in the middle of vocalization, auditory cortical responses to pitch feedback perturbations triggered by the computer during vocalization were not suppressed but were enhanced as compared with listening [12-14]. These contrasting findings suggest that the neural processes resulting in suppression or enhancement of brain responses to self-produced vocalizations compared with passive listening may depend on the relative timing between the onset of vocal production and its auditory feedback. The same phenomenon may also account for suppressed brain responses to self-triggered stimulation relative to externally triggered stimulation. Unfortunately, how auditory ERP responses reflect the neural processing of the self-triggered stimuli (e.g. button press) versus externally triggered (e.g. computer control) stimuli presented during either vocalization or listening is poorly understood.

This study, thus, was designed to address whether self-initiation of a motor act has an impact on the neural processing of auditory feedback in the regulation of voice F_0 by using the frequency perturbation technique [15]. The neural responses to −200 cents voice pitch feedback perturbation were examined in two conditions where the stimulus was delivered with a delay of 500–1000 ms after the participants pressed a button or in response to computer-triggered stimuli with a 500–1000 ms inter-stimulus intervals during active vocalization and passive listening to the playback of the preproduced voice. Our hypothesis was that engaging the motor system for self-initiation of voice F_0 feedback perturbations would enhance the processing of incoming sensory feedback, and consequently, would result in differential online processing of self-triggered and externally triggered pitch perturbations during both vocalization and listening.

Materials and methods

Participants

Twelve students (10 female and two male students, 18–27 years old) from Northwestern University, who are native-English speakers, participated in the experiments. They were all right-handed, and reported no history of speech, hearing, neurological disorders, or voice training. All participants signed consent forms approved by the North-western University Institutional Review Boards.

Apparatus

Participants were seated in a sound booth and wore an AKG boomset microphone (C420) (Harman International Industries Inc., Stamford, Connecticut, USA). Etymotic earphones (ER1-14A) (Etymotic Research Inc., Elk Grove Village, Illinois, USA) were inserted bilaterally into the participant's ear canals. The gain of the acoustical feedback pathway was calibrated with a Zwisklocki coupler and a Brüel & Kjaer sound level meter (model 2250) to a gain of 10 dB SPL between the voice level and the ear canal before testing. For the listening tasks, the amplitude of the signal to be presented to the participants was physically calibrated to be the same as that which the participants heard during their vocalizations. The vocal signal was amplified with a Mackie mixer (model 1202) (Mackie mixer, LOUD Technologies Inc., Woodinville, Washington, USA), and shifted in pitch with an Eventide Eclipse Harmonizer (Eventide Inc., Little Ferry, New Jersey, USA), and then amplified with a Crown D75 amplifier (Crown Audio Inc., Elkhart, Indiana, USA) and HP 350 dB (PreSonus Audio Electronics Inc., Baton Rouge, Louisiana, USA) attenuators. Max/MSP (Cycling 74, San Francisco, California, USA)) was used to control the harmonizer and generate a TTL pulse to mark the onset and offset of each pitch-shift stimulus (PSS) for the analysis of ERPs.

A cap consisting of Ag–AgCl electrodes (10–20 system) was used to record the Cz EEG signals (reference, linked earlobes, forehead ground). EEG signals were amplified at 10 kHz with a Grass P511 AC amplifier (Astro-Med Inc., West Warwick, Rhode Island, USA), sampled at 10 kHz using PowerLab A/D Converter (ML880, AD Instruments, Castle Hill, Australia) together with voice, feedback, TTL pulses, and recorded on a Macintosh computer by Chart software (AD Instruments, Castle Hill, Australia). The impedance of the electrode at Cz was measured using a Grass impedance meter (EZM-5AB) and maintained below 5 kΩ.

Procedure

The experiment consisted of the following four tasks: motor-vocal, motor-auditory, vocal-only, and auditory-only tasks. In the motor-vocal task, the participants were instructed to vocalize a vowel sound (/a/, 2 s duration) while holding their face, jaw, and tongue in a stationary position. In the middle of the vocalization, they were also instructed to click a wireless mouse one time to trigger a PSS (−200 cents, 200 ms duration). By instructing the participants to vocalize without contracting speech muscles, we greatly reduced the chance that such muscle contraction would affect the quality of the EEG signals. Moreover, throughout testing, we carefully monitored the EEG signals to make sure that there were no large amplitude artifacts during vocalization. After the mouse click, there was a variable delay of 500–1000 ms before the onset of the PSS. The participants were instructed to take a pause for approximately 3–4 s before the next vocalization to avoid vocal fatigue and were asked to vocalize 100 times during this task. The motor-auditory task was performed by participants pressing the mouse to trigger the PSS (−200 cents; 500–1000 ms delay after button press) during 100 playbacks of a single pre-recorded vocalization of the self-produced vowel sound /a/ (3–4 s duration). During this task, the time intervals between playback of the vocalizations were controlled by the computer at 3.5 s. In the vocal-only task, participants were instructed to vocalize the vowel /a/ for approximately 5 s during which their voice feedback was automatically shifted −200 cents (200 ms) five times by the computer with a randomized interstimulus interval of 500–1000 ms. They were asked to vocalize 30 times, yielding 150 trials for each participant. The pitch-shifted feedback channel for this task was recorded and converted to a sound file. In the auditory-only task, this sound file was then played back and the participants were asked to do nothing but passively listen to the sound.

Data analysis

The recorded signals were analyzed in IGOR Pro (v.6.0, Wavemetrics, Inc., Lake Oswego, Oregon, USA) using an event-related averaging technique. EEG signals were filtered off-line with a band-pass filter (1–30 Hz) and were then cut into epochs ranging from –200 to 700 ms, relative to the onset of the stimulus. Artifact rejection was performed before averaging, where the trials with amplitudes exceeding $\pm 50 \, \mu V$ were rejected for further data analysis. The baseline of the remaining epochs was corrected, and the ERPs were obtained for all conditions, time-locked to the stimulus onset separately for each condition. The latency and amplitude of the N1 and P2 components were extracted from the averaged neural responses by finding the peaks in 50-ms-long time windows centered at 100 and 200 ms. Repeated-measures analyses of variance were computed to test the significant differences across the tasks using SPSS (v. 16.0) (SPSS Inc., Chicago, Illinois, USA).

Results

Figure 1 shows the grand averaged auditory responses (Cz) across all the tasks. The solid lines represent the auditory responses to pitch feedback perturbations for the motor-vocal and the motor-auditory tasks and the dashed lines represent the vocal-only and the auditory-only tasks. Figure 2 shows the box plots of auditory ERP responses of N1 and P2 components across all the tasks. As can be seen, greater response magnitudes were associated with motor-vocal and motor-auditory tasks as compared with vocal-only and auditory-only tasks, respectively.

Repeated-measures analyses of variance were performed to test the effect of self-triggered motor activity on the modulation of auditory ERPs during vocalization and listening, respectively. During vocalization, both N1 and P2 response magnitudes were significantly larger for the motor-vocal task than those for the vocal-only task [N1: $F(1,11) = 20.006$, $P = 0.001$; P2: $F(1,11) = 28.619$, $P < 0.001$]. Similarly, during listening, both N1 and P2 responses showed larger magnitudes for the motor-auditory task than those for the auditory-only task [N1: $F(1,11) = 18.054$, $P = 0.001$; P2: $F(1,11) = 39.300$, $P < 0.001$]. In addition, the latencies of both N1 and P2 responses did not differ between motor-vocal and vocal-only task or between motor-auditory and auditory-only task.

Discussion

This study was designed to address how a self-triggered nonvocal motor act affects the online processing of voice pitch auditory feedback perturbation during vocalization and while listening to the playback of self-produced vocalizations. The results showed that nonvocal motor acts (motor-vocal and motor-auditory) elicited larger-amplitude N1 and P2 responses than vocal-only and auditory-only tasks, respectively. These findings indicated that auditory responses to self-triggered pitch perturbations were enhanced relative to those triggered externally.

Earlier research has shown that neural responses to a self-triggered tone were suppressed when compared with those triggered externally [3,4,6,7,16]. The suppression of neural responses occurs when a pure tone follows a self-generated motor act by a fixed time [6], and it is reduced when there was a variable delay between a button press and the presentation of a tone [7]. However, it has also been shown that ERP responses to pitch-shifted feedback during ongoing vocalization were enhanced compared with responses triggered by stimuli delivered at the onset of vocalization [12,14]. In this study, pitch-shifted stimuli during vocalization were either delivered by a 'mouse click' (motor-vocal) or triggered externally by a computer (vocal-only), and it was found that the 'mouse click' led to enhanced responses during both the vocalizing and listening conditions compared with

computer-triggered conditions. This finding of motor enhancement of neural responses is similar to the earlier study reported by Behroozmand *et al.* [14].

It is speculated that procedural differences between this study and earlier research may explain the differences in the neural responses reported here. We used a nonvocal motor task to trigger pitch-shifted stimuli during vocalization as an attempt to combine self-triggered stimuli and pitch-shifted voice feedback to define the variables leading to the different responses. The results showed enhanced responses to pitch-shifted voice feedback triggered by a nonvocal motor act during vocalization. For this case, the key variable differentiating these studies appears to be vocalization itself; specifically, the efference copy of the predicted vocal output. Thus, it is suggested that the cortical processing of voice pitch feedback perturbation may be enhanced by motor cortical activity involved in the mouse clicks during vocalization. This effect may be one potential explanation for the motor-induced enhancement of neural responses to self-triggered stimulation.

However, it was also observed that perturbations of the participant's prerecorded vocalization, which were triggered by a nonvocal motor act also led to enhanced responses compared with passive listening to the vocalizations with perturbations. Activation of the vocal motor cortex cannot explain these results. The most obvious explanation of these findings is that the participant recognized the voice as belonging to 'self'. According to previous studies in which participants heard tones that were self-triggered by a button press, responses were suppressed relative to externally triggered tones. The only apparent difference between the latter studies and the motor-auditory and auditory-only tasks of this study was the recognition of self-vocalization. Therefore, it appears that a nonvocal motor act that triggers a perturbation to one's own prerecorded vocalization may enable cortical mechanisms that identify self-vocalization. In this case, a nonvocal motor act may facilitate the responsiveness of auditory cortex to perturbations in self-vocalization to a greater degree than the suppressive effects induced by motor activity, leading to the larger neural responses in the motor-auditory task than those in the auditory-only task.

It may be argued that a motor-only task should have been included in this study, where participants would press the button in a similar manner as in the motor-vocal or motor-auditory tasks but no pitch feedback perturbation would be presented. This task would be used to rule out the possibility that enhancement of N1 and P2 responses in the motor-vocal or motor-auditory tasks could be because of the contamination of brain signals related to hand movement. However, we think this is a remote possibility for the following reasons: the primary fluctuations of brain responses caused by a motor act occur within 400 ms with respect to the onset of hand movement, and can be ignored after this period [4,7]. In this study, stimuli were delivered with a 500–1000 ms delay after the mouse click, and brain responses were measured with respect to the onset of stimulus delivery rather than the onset of mouse click. Thus, under such circumstances, the contamination of motor activity can be ruled out and should have had no effect on the present findings.

It is noteworthy that one confounding factor that may possibly affect our findings is the number of stimuli across the tasks: five PSS were presented during each vocalization for vocal-only and auditory-only tasks, whereas only one for motor-vocal and motor-auditory tasks. In review of earlier studies, however, we think that this confounding factor can be ruled out. First, behavioral results showed that vocal responses to pitch feedback perturbations did not differ between a 1 PSS condition and 5 PSS condition during each vocalization [17]. In contrast, findings from our ERP studies showed a constant vocalization-induced enhancement when pitch feedback was shifted eight times [12] or one time [14]. Therefore, these previous findings indicate that the factor of the number of PSS across the tasks would have not influenced the validity of our conclusion.

Conclusion

This study showed that auditory ERPs to self-triggered pitch perturbation were enhanced when compared with externally triggered stimuli presented in the middle of both vocalization and listening. These observations suggest that vocalization or the recognition of one's own voice facilitates the responsiveness of the auditory cortex to perturbations in auditory feedback.

Acknowledgments

References

1. Blakemore SJ, Wolpert DM, Frith CD. Central cancellation of self-produced tickle sensation. Nat Neurosci. 1998; 1:635–640. [PubMed: 10196573]

2. Blakemore SJ, Wolpert D, Frith C. Why can't you tickle yourself? Neuroreport. 2000; 11:R11–R16. [PubMed: 10943682]

3. Schafer EW, Marcus MM. Self-stimulation alters human sensory brain responses. Science. 1973; 181:175–177. [PubMed: 4711735]

4. Martikainen MH, Kaneko K, Hari R. Suppressed responses to self-triggered sounds in the human auditory cortex. Cereb Cortex. 2005; 15:299–302. [PubMed: 15238430]

5. Wolpert DM. Computational approaches to motor control. Trends Cogn Sci. 1997; 1:209–216. [PubMed: 21223909]

6. Aliu SO, Houde JF, Nagarajan SS. Motor-induced suppression of the auditory cortex. J Cogn Neurosci. 2009; 21:791–802. [PubMed: 18593265]

7. Bass P, Jacobsen T, Schroger E. Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. Int J Psychophysiol. 2008; 70:137–143. [PubMed: 18627782]

8. Heinks-Maldonado TH, Nagarajan SS, Houde JF. Magnetoencephalographic evidence for a precise forward model in speech production. Neuroreport. 2006; 17:1375–1379. [PubMed: 16932142]

9. 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]

10. Ventura MI, Nagarajan SS, Houde JF. Speech target modulates speaking induced suppression in auditory cortex. BMC Neurosci. 2009; 10:58. [PubMed: 19523234]

11. Curio G, Neuloh G, Numminen J, Jousmaki V, Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. Hum Brain Mapp. 2000; 9:183–191. [PubMed: 10770228]

12. 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]

13. Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature. 2008; 453:1102–1106. [PubMed: 18454135]

14. Behroozmand R, Liu H, Larson CR. Time-dependent neural processing of the auditory feedback during voice pitch error perturbation. J Cogn Neurosci. in press.

15. 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]

16. McCarthy G, Donchin E. The effects of temporal and event uncertainty in determining the waveforms of the auditory event related potential (ERP). Psychophysiology. 1976; 13:581–590. [PubMed: 996223]

17. Bauer JJ, Larson CR. Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: improvements in methodology for the pitch-shifting technique. J Acoust Soc Am. 2003; 114:1048–1054. [PubMed: 12942983]
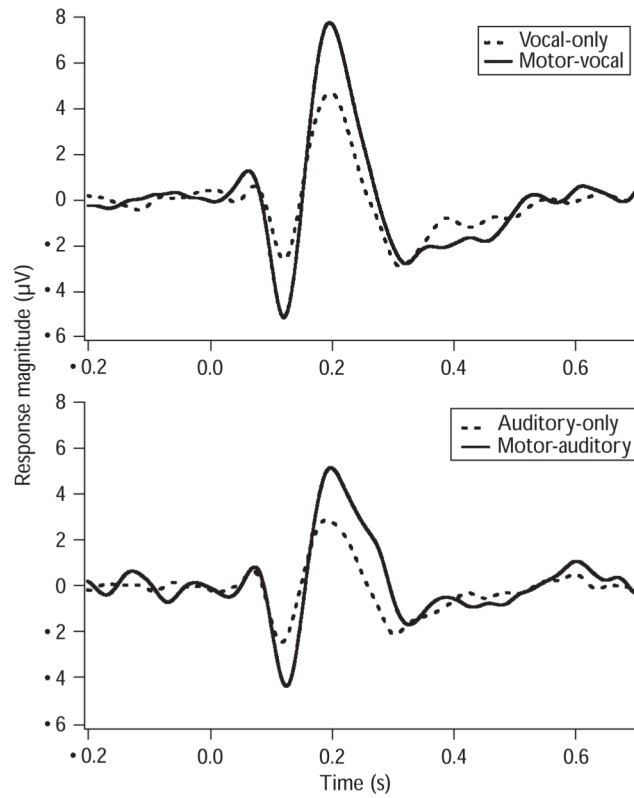
Fig. 1.
Auditory event-related potential waveforms (Cz) averaged over all participants across the conditions. The solid line denotes the auditory responses for the motor-vocal task and the motor-auditory task, and the dashed line for the vocal-only and the auditory-only tasks. Zero at horizontal axis represents where the pitch perturbations begin.
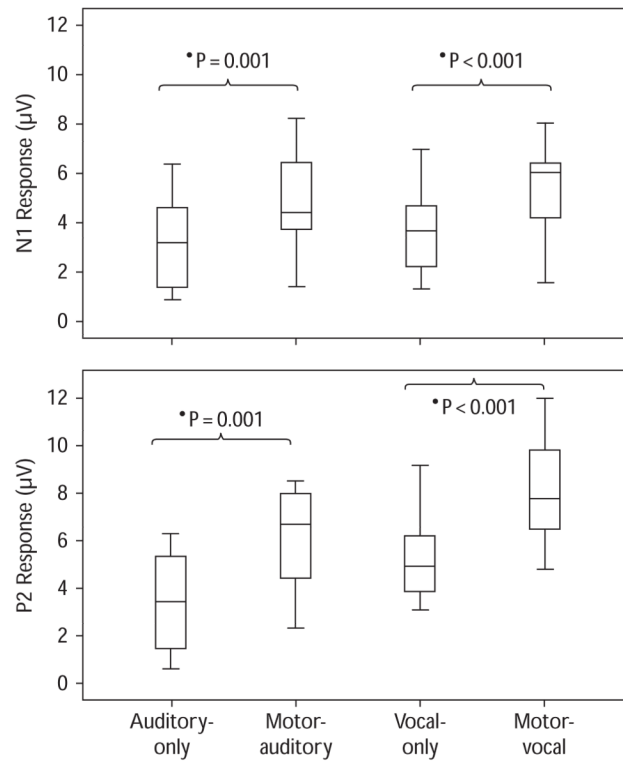
Fig. 2.
Box plots of N1 (top) and P2 (bottom) response magnitudes across all the tasks. Box plot definitions: middle line is median, top and bottom of boxes are 75th and 25th percentiles, whiskers extend to limits of main body of data defined as high hinge + 1.5 (high hinge – low hinge) and low hinge –1.5 (high hinge – low hinge).