2016

Neural And Behavioral Correlates of Planning During Vocal Motor Control

Stacey Jean Sangtian
University of South Carolina

Recommended Citation

Sangtian, S. J.(2016). *Neural And Behavioral Correlates of Planning During Vocal Motor Control.* (Master's thesis). Retrieved from https://scholarcommons.sc.edu/etd/3777

NEURAL AND BEHAVIORAL CORRELATES OF PLANNING DURING VOCAL MOTOR CONTROL

by

Stacey Jean Sangtian

Bachelor of Science
College of Charleston, 2011

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Speech Pathology in

Speech Pathology

The Norman J. Arnold School of Public Health

University of South Carolina

2016

Accepted by:

Roozbeh Behroozmand, Director of Thesis

Allen Montgomery, Reader

Dirk-Bart den Ouden, Reader

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

ACKNOWLEDGEMENTS

I would like to express my appreciation to my thesis director and research mentor, Dr. Roozbeh Behroozmand, for his unwavering support, encouragement, and guidance throughout the completion of this project. He has granted me every opportunity to develop my skills as an independent researcher and without him this thesis would not have been possible. In addition, I would like to thank Dr. Allen Montgomery and Dr. Dirk-Bart den Ouden for taking the time to serve on my thesis committee. Their positive support of my research pursuits and their thoughtful input were greatly appreciated.

I would also like to thank Karim Johari for his assistance with subject preparation; my classmates LCL, LM, and SA for their constant support and encouragement throughout this process; the clinical and academic faculty of the Department of Communication Sciences and Disorders for their support of my research goals; and the subjects who gave their time to participate in this study.

Finally, I would like to thank my parents for their encouragement and support of all the extra projects I took on over the past two years. I would not have been able to accomplish everything without them.

ABSTRACT

Speech is a goal-directed movement developed to serve the purpose of human communication. During speech, the brain simultaneously coordinates complex motor behavior and sensory information for planning, execution and online monitoring for error correction in order to ensure accurate delivery of the vocal messages. Models of speech production and voice motor control have adapted the principles of the internal forward model to explain that speech sound production is controlled by comparison of internal predictions with sensory feedback. Growing evidence from studies conducted in speech and limb motor movement indicate that these sensory predictions and responsive commands are loaded in anticipation of voluntary movement. In the present study, we investigated the generation of predictive sensorimotor plans during the planning phase of speech production and the influence of these plans on voice motor control. We hypothesized that the internal forward model is active, and the efference copies are present, during the planning phase of vocal motor movement. Further, we hypothesized that neurophysiological and behavioral mechanisms of vocal motor control are modulated during the planning phase of vocal motor movement. Subjects were instructed to maintain a steady vocalization of a vowel sound and change their voice pitch in upward and downward directions following the onset of a visual cue. During the planning phase of sustained vocalizations, upward and downward pitch-shift stimuli randomly perturbed the speech auditory feedback. Behavioral results revealed compensatory vocal responses and differences in onset latency time and peak magnitude for conditions that loaded a motor plan (i.e., cue up or cue down). Neurophysiological responses revealed direction-specific modulation of P1 responses, N1 suppression for conditions that loaded a motor plan regardless of direction, and modulation of the P2 component only for the upward stimulus direction. These findings contribute to the understanding of the mechanism underlying motor planning during preparation for voice motor control.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The ability to produce movement is a key function that subserves many different aspects of behavior in animals and humans. A large category of movements, termed goal-directed movements, are generated to reach a target to accomplish the goals of a behaviorally-relevant task (e.g., grabbing a cup of coffee or hitting a tennis ball). The question of how the brain controls goal-directed movements has been debated for decades. A widely-accepted answer to this question has been provided by the internal forward model theory (Wolpert and Flanagan, 2001), which hypothesizes that goal-directed movements are controlled by a mechanism that provides internal predictions about the sensory consequences of self-generated motor actions. This theory explains that the internal forward model is part of a predictive coding mechanism in which the efference copies of the motor commands are translated into internally-predicted sensory representation of self-generated movements (Wang et al., 2014; Wolpert et al., 1995). By incorporating the actual sensory feedback and comparing it with the internal predictions, the brain can estimate prediction errors and use them to correct subsequent motor behavior during execution and control of goal-directed movements.

Speaking is one type of goal-directed movement developed to serve the purpose of human communication. During speech, the brain has to coordinate a highly complex motor behavior involving the regulation of muscle movements in the articulatory, phonatory, and respiratory systems for producing sounds that can be perceived by other listeners. At the same time, sensory information in the form of somatosensory (Lametti et al., 2012) and auditory feedback (Cai et al., 2011) is being processed by the brain for online speech monitoring and error correction in order to ensure accurate delivery of the vocal messages. The recent models of speech production have adapted the principles of the internal forward model to explain how speech sound production is controlled by incorporating sensory feedback and comparing it with internal predictions (Guenther, 2006; Guenther et al., 2006; Hickock et al., 2011; Houde and Chang, 2015; Houde and Nagarajan, 2011).

Studies on the mechanisms of voice motor control have provided supporting evidence for the internal forward model theory particularly with increased attention to the role of auditory feedback in the processing of speech (Houde and Chang, 2015). To further examine the role of auditory feedback, studies have shown that applying pitch shift stimuli to the auditory feedback during the vocal task of sustained vowel phonations elicits behavioral and neurophysiological event-related potential (ERP) responses that highlight important aspects of sensory-motor mechanisms involved in voice motor control (Behroozmand et al., 2009; Behroozmand and Larson, 2011; Z. Chen et al., 2012; Heinks-Maldonado et al., 2005; Heinks-Maldonado et al., 2006; Houde et al., 2002; Korzyukov et al., 2012; Scheerer and Jones, 2014). Behavioral responses to direction and magnitude of pitch shifts in voice auditory feedback have revealed two types of responses: opposing and following (Burnett et al., 1998). It has been suggested that opposing responses to pitch shift are reflexive attempts to correct and compensate for the internal mismatch between the efference copy and auditory feedback to stabilize voice fundamental frequency (F0). Conversely, it has been proposed that the following vocal responses to pitch shift occur due to the perturbation being perceived as an external referent that is to be matched (Behroozmand et al., 2012). The neurophysiological correlates of voice motor control mechanisms have also been investigated by examining the amplitude changes of the P1-N1-P2 ERP components in response to pitch shifts in voice auditory feedback (Behroozmand et al., 2009; Korzyukov et al., 2012; Scheerer et al., 2013). For example, a study by Korzyukov et al. (2012) found that predictable direction (upward or downward) of pitch perturbation resulted in less opposing (compensatory) vocal responses and reduced amplitude of N1 as compared to unpredictable changes in direction. This study, among others (e.g., Behroozmand et al., 2012; Scheerer and Jones, 2014), have provided evidence that predictions about auditory stimuli modulates compensatory behaviors during a vocal motor task.

While the above studies examined modulation of responses after initiation of a motor task, growing evidence from studies conducted in speech (Daliri and Max, 2015; Max et al., 2008) and limb motor movement (Ahmadi-Pajouh et al., 2012; Carlsen et al., 2012) indicate that sensory predictions and responsive commands are loaded in anticipation of voluntary movement, i.e., during the planning phase before voluntary movement initiation. In the speech domain, Daliri and Max (2015) and Max et al. (2008) administered tones during the planning phase of speaking, reading and seeing conditions. They found that attenuation of the N1 ERP component occurred during the movement planning phase in the speaking condition but not in the planning phase of the silent reading or seeing conditions. Their findings indicated that processing of sensory components relevant to the anticipated speech production occurred during the planning phase of the speech motor task. Additionally, studies on limb motor movement have shown that presentation of a startling acoustic stimulus during the planning phase of a voluntary limb motor task evokes release of the intended movement (see Carlsen et al. (2012) for a review) further supporting the notion of pre-programmed controls for movement.

Perturbations occurring during this planning phase have also been shown to modulate compensatory behaviors. One such work supporting this notion is a study by Ahmadi-Pajouh et al. (2012) on limb motor control in which subjects adapted to a force field on a reach task and received an external perturbation during the planning phase of the task that displaced the limb in the same or opposite direction of the anticipated force field. Measurements of hand displacement and electromyography (EMG) showed that after adaptation, perturbations that displaced the hand in the same direction as the expected force field during the planning phase of a reach resulted in smaller hand displacement and an increase in feedback gains in the long-latency period (45-100 ms delay) as compared to baseline. These results indicated that feedforward commands and feedback gains specific to the anticipated movement are pre-programmed during the planning phase of a motor task thereby affecting behavioral responses to disruptions during this period.

In the present study, we aimed to investigate the neural and behavioral mechanisms of planning during a vocal motor task utilizing the altered auditory feedback paradigm to measure vocal behavior and ERP responses. The purpose of this study was to answer the following questions: Is the internal forward model active, and the efference copies present, during the planning phase of vocal motor movement? Are the mechanisms of vocal motor control modulated during the planning phase of vocal motor movement? In order to address these questions, we designed an experiment in which subjects maintained steady vocalizations of the vowel sound /a/ and were prompted by a visual cue to prepare for performing one of the following voluntary vocal pitch modulation tasks: 1) up arrow: raise the voice pitch, 2) down arrow: lower the voice pitch, and 3) dash line: maintain the same voice pitch (no change: control). Before subjects voluntarily modulated their voice pitch, an upward or downward pitch-shift stimulus (PSS, 100 cents) was randomly delivered to perturb the auditory feedback of their steady vowel vocalization during the planning phase.

Based on our understanding of the internal forward model theory in speech production (Guenther, 2006; Guenther et al., 2006; Hickock et al., 2011; Houde and Chang, 2015; Houde and Nagarajan, 2011) and current evidence that supports pre-programming of sensory predictions in anticipation of a goal-oriented movement (Ahmadi-Pajouh et al., 2012; Daliri and Max, 2015; Max et al., 2008), we hypothesized that the internal forward model is active, and the efference copies are present, during the planning phase of vocal motor movement. Further, we hypothesized that mechanisms of vocal motor control are modulated during the planning phase of vocal motor movement. This would be evidenced by modulation of neurophysiological (i.e., ERP) and behavioral responses to stimuli. This study provides insights to the effect of motor planning on control of action in the speech domain and contributes to the understanding of the underlying neural and behavioral mechanisms of vocal responses to perturbations in auditory feedback.

CHAPTER 2

MATERIALS AND METHODS

2.1 SUBJECTS

Data for this study was collected from 18 subjects (7 male and 11 female, age range: 18-29 years, mean age: 22.95 years). The subjects were right-handed, native speakers of English with normal hearing and no history of neurological or speech disorders. Subjects were either monetarily compensated or received course credit at the University of South Carolina for their participation. All experimental procedures were approved by the University of South Carolina Institutional Review Board.

2.2 EXPERIMENTAL DESIGN

Subjects were seated in a sound attenuated booth in which EEG signals and vowel sound vocalizations were recorded. The experimental task consisted of two parts: the planning phase and the voluntary pitch modulation vocal task (Figure 2.1). During the planning phase, black-colored directional cues (up arrow, down arrow, or dash line) were visually presented in random order and subjects were instructed to begin vocalizing a steady /a/ vowel at their conversational pitch and loudness and prepare to perform one of the following vocal tasks after the visual cue turned green (go signal): 1) up arrow: raise the voice pitch, 2) down arrow: lower the voice pitch, and 3) dash line: maintain the same pitch (no change: control). In order to perform the vocal tasks, subjects were instructed to smoothly glide their voice pitch in the direction of the arrow following the green color change of the cue. For the dash line (control condition), the color of the visual cue did not

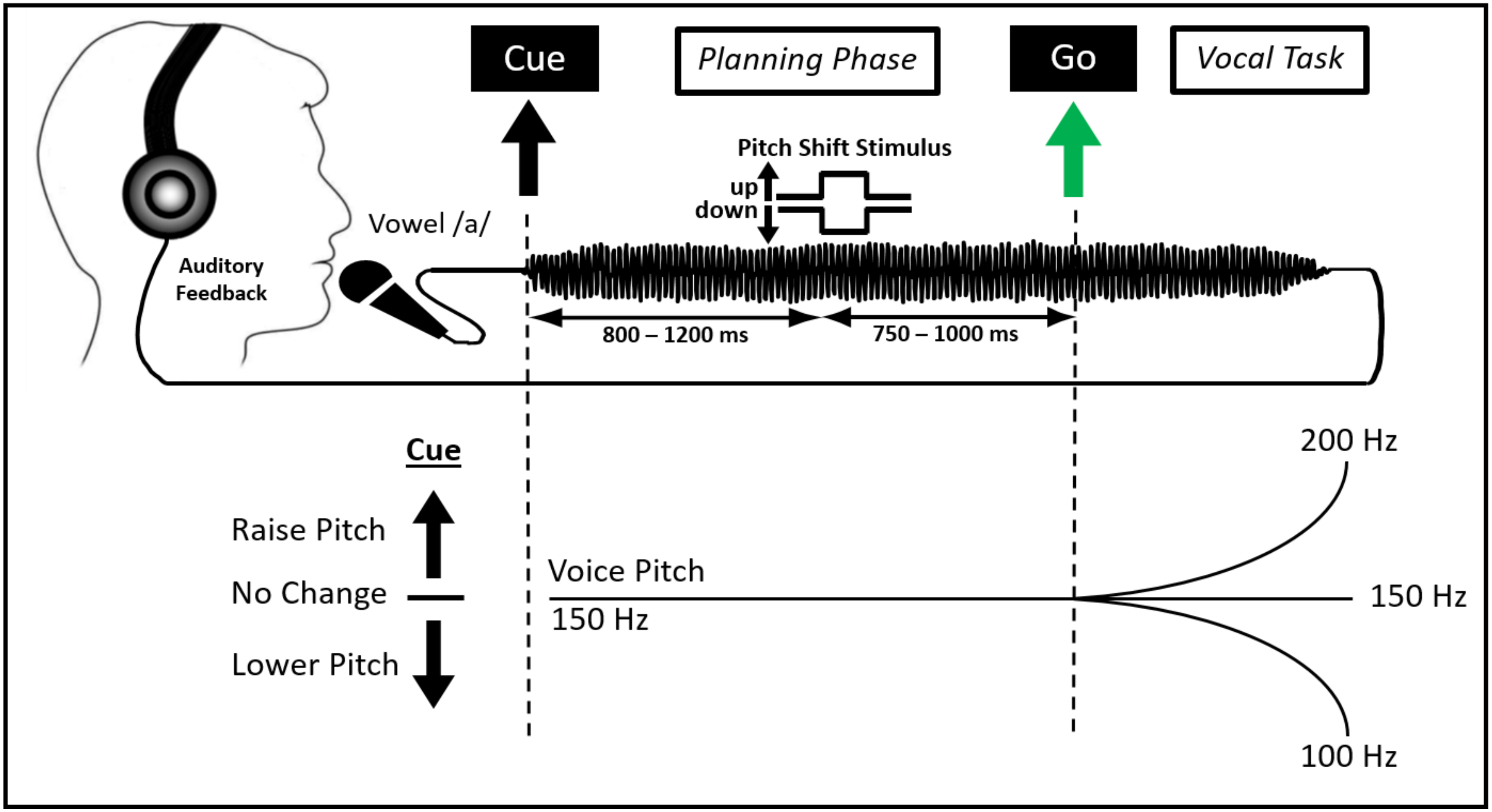

Figure 2.1 Experimental Design Directional cues (black arrow or horizontal dash) and pitch shift direction (+/- 100 cents) were randomized. Dashed vertical lines indicated onset of steady pitch phonation and onset of pitch glides. Pitch shift stimulus (PSS) onset time was randomized between 800-1200 ms after onset of the steady pitch vocalization. “Go” cue (green arrow) was randomized between 750-1000 ms following PSS onset.

change to green, indicating that there was no need to change the voice pitch during steady vowel sound vocalizations. During each vocalization trial, a brief (200 ms) pitch shift stimulus perturbed voice auditory feedback in the planning phase with onset latencies randomized between 800-1200 ms after the onset of the vocalization. The direction of stimuli was randomized between upward (+100 cent) and downward (-100 cents) pitch-shifts across trials. The vocal task was initiated by onset of the green arrows (go signal) that appeared at a randomized time between 750-1000 ms relative to the onset of the pitch-shift stimulus during the planning phase.

The combination of the pitch-shift stimulus and the vocal task cues led to a total of 6 different conditions in this study (an upward or downward pitch shift in voice feedback for voluntary voice changes in upward or downward directions or no change). Short (2 – 3 s) breaks were given between successive trials. A total number of 400 vocalizations (about 66 vocalizations per condition) were produced and recorded during each session.

2.3 VOICE AND EEG DATA ACQUISITION

The voice data was picked up using a head-mounted AKG condenser microphone (model C520), amplified by a Motu Ultralite-MK3 and recorded at 44.1 kHz on a laboratory computer. A custom-designed program in Max/Msp (Cycling 74, v.5.0) controlled an Eventide Eclipse Harmonizer to pitch shift the voice online and feed it back to the ears using Etymotic earphones (model ER1-14A). The Max/Msp program also controlled all aspects of the visual cues and stimuli (e.g. direction, onset time etc.) and generated TTL pulses to accurately mark the onset of each event during preparatory and vocal task periods for all trials. A 10 dB gain between voice and its feedback was maintained to partially mask air-born and bone-conducted voice feedback during vocalizations.

The EEG signals were recorded from 64 sites on the subject's scalp using an Ag–AgCl electrode cap (Easy-Cap GmbH, Germany) with an average reference montage. A BrainVision actiCHamp amplifier (Brain Products GmbH, Germany) on a computer utilizing Pycorder software recorded the EEG signals at a 1 KHz sampling rate after applying a low-pass anti-aliasing filter with a 200 Hz cut-off frequency.

2.4 ANALYSIS OF BEHAVIORAL VOCAL RESPONSES

The pitch frequency of the recorded voice signals was extracted in Praat (Boersma and Weenik, 1996) using an autocorrelation method and then exported to MATLAB for further processing. The extracted pitch frequencies were segmented into epochs ranging from −100 ms before to 500 ms after the onset of pitch-shift stimuli. Pitch frequencies were converted from Hertz to the Cents scale to calculate vocal compensation in response to the pitch-shift stimulus using the following formula:

$$\textit{Vocal Compensation [Cents]} = 1200 \times \log_2(F/F_{\text{Baseline}})$$

F was the post-stimulus pitch frequency and F_{Baseline} was the baseline pitch frequency from −100 to 0 ms pre-stimulus. The calculated pitch contours in Cents was averaged across all trials in each of the 6 experimental conditions separately. The extracted pitch contours were then averaged across all subjects to obtain the grand-averaged profile of the vocal responses to pitch-shift stimulus for each condition. The onset latency of the vocal responses was calculated as the time of the first point at which the magnitude of the responses exceeded the baseline by ±2 standard deviations. The vocal response peak magnitudes were extracted for the first prominent peak in a time window from 0–500 ms post-stimulus.

2.5 EEG DATA ANALYSIS

The EEGLAB toolbox (Delorme and Makeig, 2004) was used to analyze recorded EEG signals in order to calculate ERPs time-locked to the onset of upward or downward pitch-shift stimuli during the planning phase of the task. Recorded EEG were filtered offline using a bandpass filter with cut-off frequencies set to 1 and 30 Hz (−24 dB/oct) and then segmented into epochs ranging from −100 ms before and 500 ms after the onset of the stimulus. Artifact rejection was then carried out by excluding epochs with EEG amplitudes exceeding ±50 μV. Individual epochs were subjected to baseline correction by removing the mean amplitude of the pre-stimulus time window from -100 to 0 ms for each electrode. The extracted epochs were averaged across all trials separately for each condition to obtain the ERP responses to pitch shift in individual subjects. A minimum number of 45 trials were used to calculate the ERP responses for each subject. The extracted ERP profiles were then averaged across all subjects to calculate the grand-averaged ERP responses and the amplitude of P1-N1-P2 components were extracted within a time window centered around latencies calculated based on the peak amplitude of the ERP components at the Cz electrode (vertex).

2.6 STATISTICAL ANALYSIS

A 2×3 repeated-measures analysis of variance (ANOVA) was employed to investigate main effects of pitch-shift stimulus direction (up vs. down) and vocal task (raise pitch, lower pitch or no change), and their interactions on the latency and magnitude of vocal and ERP responses to auditory feedback perturbations.

CHAPTER 3

RESULTS

3.1 BEHAVIORAL VOCAL RESPONSES

Results of the analysis for grand-average vocal responses to pitch shifted auditory feedback are shown in Figure 3.1. As shown in this figure, subjects produced a compensatory vocal response that opposed the direction of both upward (+100 cents) and downward (-100 cents) pitch-shift stimuli for cue up, cue down and control conditions. A 2×3 Rm-ANOVA on the latency of vocal responses only revealed a significant main effect of condition ($F(2,34) = 4.184$, $p<0.05$). Post-hoc tests using Bonferroni's correction for multiple comparisons revealed that the main effect of condition was accounted for by a significantly longer ($p<0.05$) latency of the vocal responses when the direction of cue matched the direction of pitch shifts for both upward and downward stimuli compared with control (Figure 3.1b and 3.1e). In addition, we also found that for upward pitch shifts, the latency of vocal responses was significantly longer for cue up vs. cue down conditions. For downward pitch shifts, the latency of vocal responses was significantly longer for cue up vs. cue down condition. A 2×3 Rm-ANOVA on the magnitude of vocal responses only revealed a significant main effect of condition ($F(2,34) = 4.994$, $p<0.05$). Post-hoc tests using Bonferroni's correction for multiple comparisons revealed that the main effect of condition was accounted for by a significant ($p<0.05$) difference between the magnitude of vocal responses to upward pitch-shift stimuli for cue up vs. cue down, as well as, cue up vs. control conditions. We found that

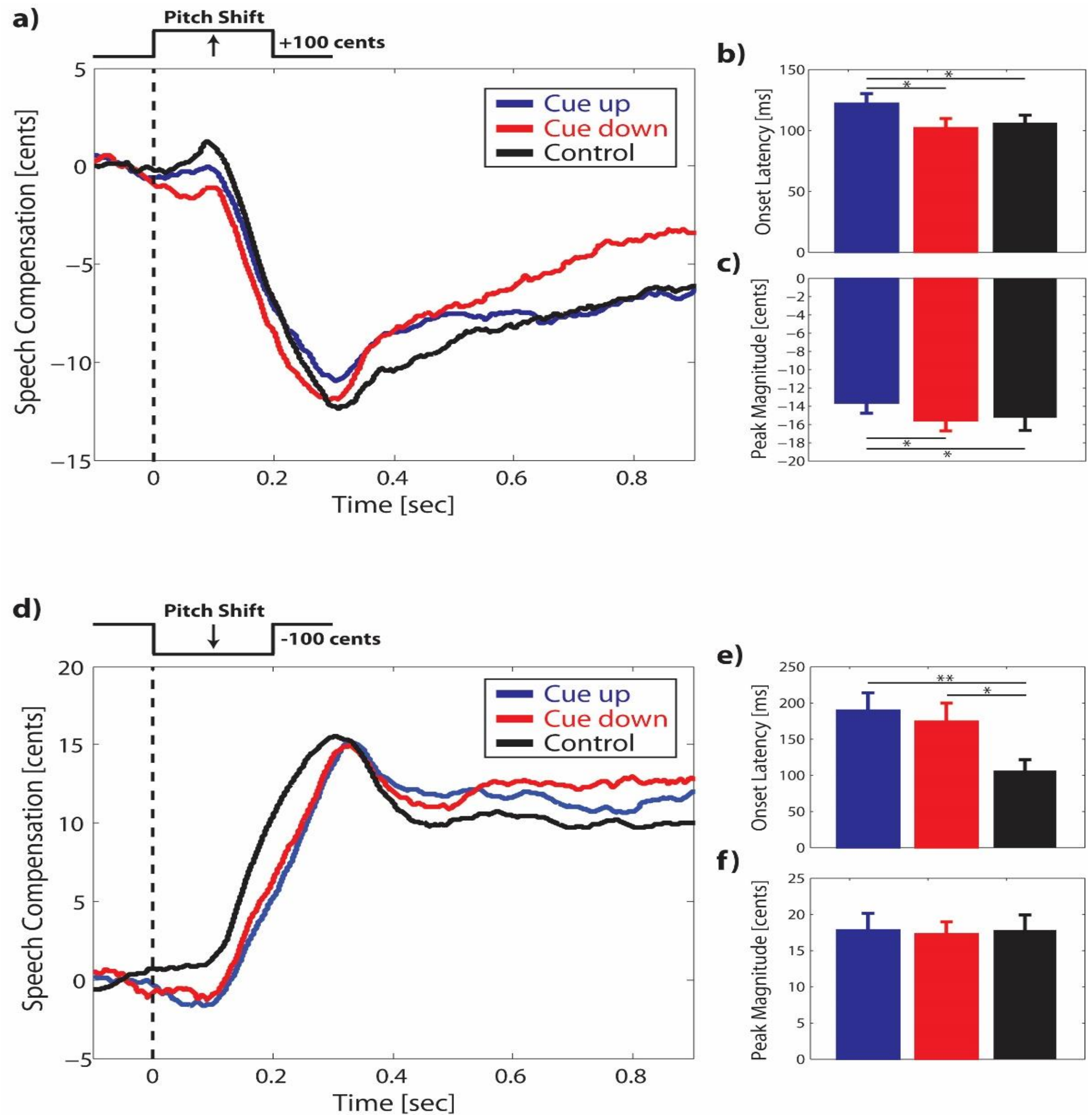

Figure 3.1 Behavioral Vocal Responses Grand-average vocal responses to pitch-shift stimulus in the auditory feedback. a) Profile of the compensatory vocal responses to upward (+100 cents) pitch-shift stimulus overlaid for cue up, cue down and control conditions. b) Comparison of the mean onset latency of vocal responses to upward stimulus in all cue conditions. c) Comparison of the mean magnitude of vocal responses to upward stimulus in all cue conditions. d) Profile of the compensatory vocal responses to downward (-100 cents) pitch-shift stimulus overlaid for cue up, cue down and control conditions. e) Comparison of the mean onset latency of vocal responses to downward stimulus in all cue conditions. f) Comparison of the mean magnitude of vocal responses to downward stimulus in all cue conditions.

the subjects produced significantly smaller compensatory vocal responses to upward pitch shifts for cue up (matching direction) vs. cue down and control conditions (non-matching direction).

3.2 ERP RESPONSES

Analysis of the grand-average ERPs led to the identification of time-locked P1, N1 and P2 components that were elicited in response to upward and downward pitch-shift stimuli at latencies approximately at 50, 100 and 200 ms (Figure 3.2). Therefore, for the analysis of the ERP data, we focused on examining the effects of stimulus direction (up vs. down) and cue (up, down and control) on the magnitude of the P1, N1 and P2 ERP components. ERP activities were subjected to statistical analysis at electrodes over the frontal (F1, Fz, F2), central (C1, Cz, C2), temporal (T7 and T8) and temporo-parietal (TP9 and TP10) areas. The choice of these electrodes was based on visual inspection of the topographical distribution of the ERP responses over the surface of the scalp (Figures 3.3-3.5).

3.3 P1 ERP COMPONENT

The P1 ERP responses were elicited at approximately 50 ms after the onset of pitch-shift stimuli in the auditory feedback during steady vocalization of the vowel sound while subjects prepared to change their vocal pitch output in the direction of the presented visual cues (up, down or control). The profile of the grand-average ERP responses to upward and downward stimuli, overlaid across cue up, cue down and control conditions are shown in Figure 3.3. The topographical distribution maps of the P1 responses showed that this component was elicited with a positive peak of activation predominantly over the frontal region and inverted (negative polarity) over the temporo-parietal region (Figure 3.3b and 3.3d). A 2×3 Rm-ANOVA on the mean amplitude of the P1 component

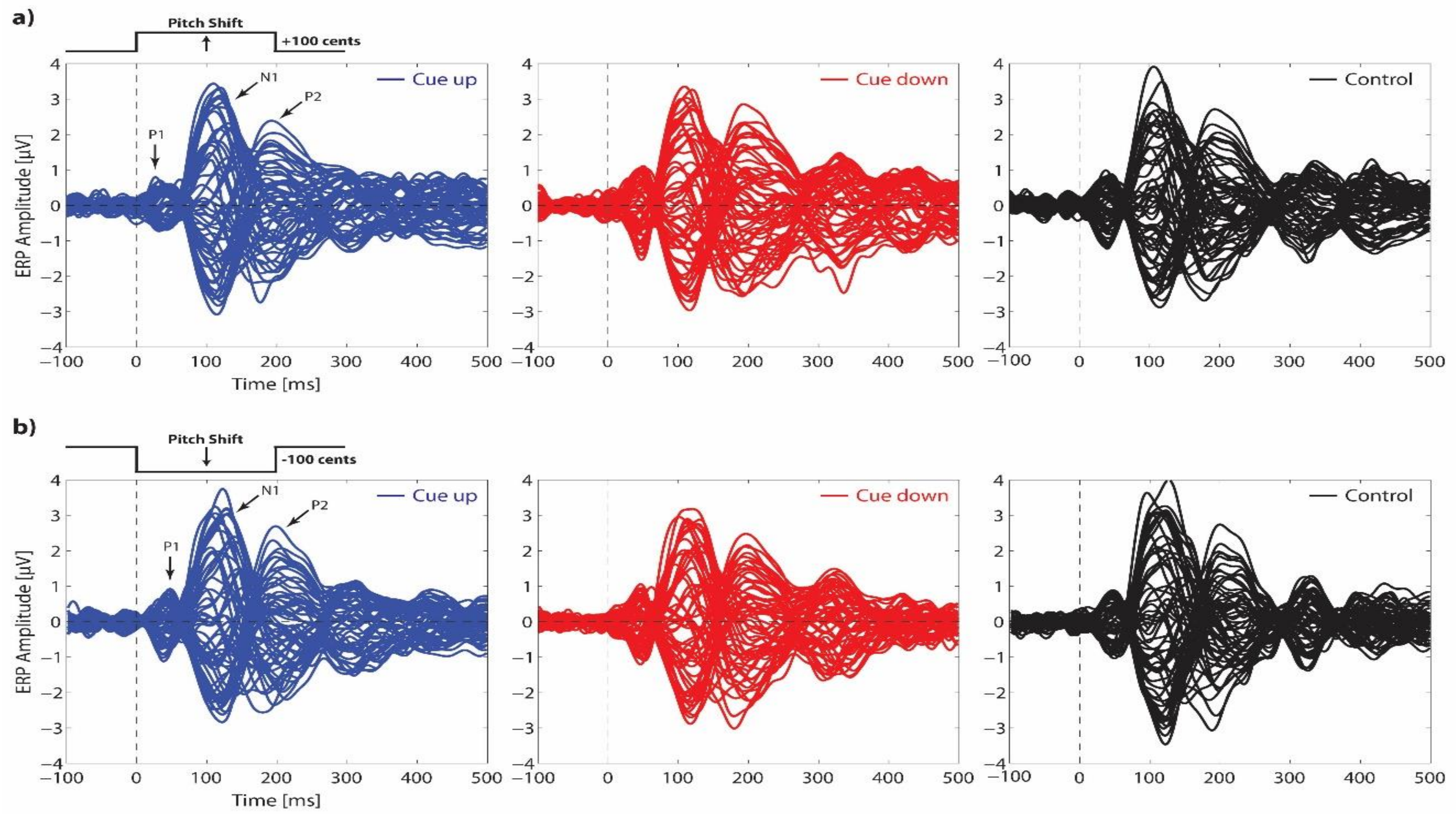

Figure 3.2 Butterfly Representation of ERP Responses The butterfly representation of the event-related potentials (ERPs) in responses to a) upward (+100 cents) and b) downward (-100 cents) pitch-shift stimuli during cue up, cue down and control conditions. In each panel, ERP responses time-locked to stimulus onset are overlaid for a total of 64 EEG recording sites on the surface of the scalp according to the standard 10-20 montage.

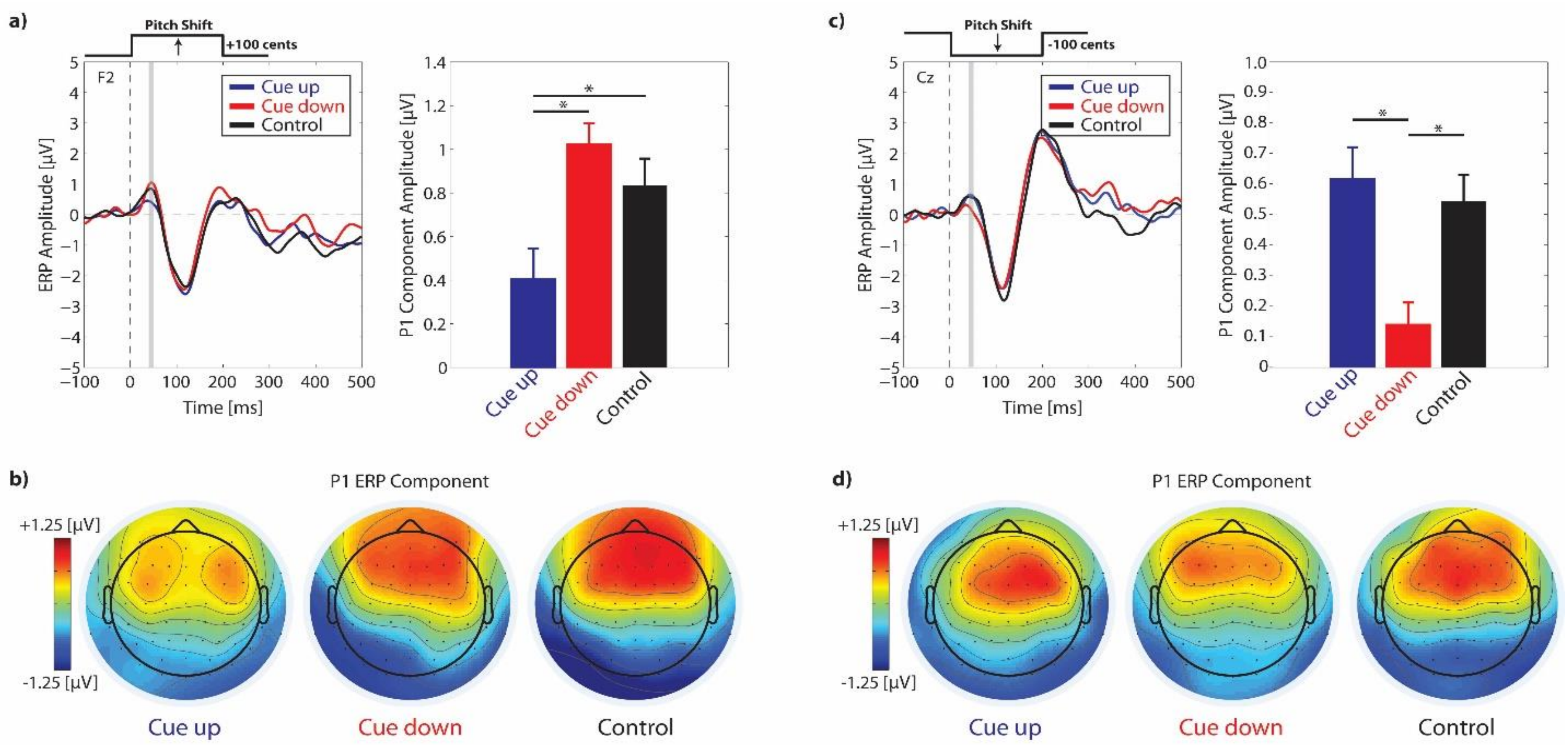

Figure 3.3 P1 ERP Response a) Profile of the overlaid grand-average P1 ERP responses to upward (+100 cents) pitch-shift stimulus for cue up, cue down and control conditions. The bar plots show the statistical comparison of the P1 component of ERPs (40-50 ms) across the three conditions. b) Topographical scalp distribution maps of the P1 component in response to upward pitch-shift stimuli for cue up, cue down and control conditions. c) Profile of the overlaid grand-average P1 ERP responses to downward (-100 cents) pitch-shift stimulus for cue up, cue down and control conditions. The bar plots show the statistical comparison of the P1 component of ERPs (40-50 ms) across the three conditions. d) Topographical scalp distribution maps of the P1 component in response to downward pitch-shift stimuli for cue up, cue down and control conditions.

from 40-50 ms only revealed a significant main effect of condition over the frontal ($F(2,34) = 4.424$, $p<0.05$) and central ($F(2,34) = 4.396$, $p<0.05$) regions. Post-hoc tests using Bonferroni's correction for multiple comparisons revealed that the main effect of condition was accounted for by a significant ($p<0.05$) suppression of P1 responses when the direction of cue matched the direction of pitch shift for both upward and downward stimuli. An example of this effect is illustrated in bar plots shown in Figures 3.3a and 3.3c.

3.4 N1 ERP COMPONENT

The N1 ERP responses were elicited at approximately 100 ms after the onset of pitch-shift stimuli in the auditory feedback. The profile of the grand-average ERP responses to upward and downward stimuli, overlaid across cue up, cue down and control conditions are shown in Figure 3.4. The topographical distribution maps of the N1 responses showed that this component was elicited with a positive peak of activation predominantly over the fronto-central region and inverted (negative polarity) over the temporo-parietal region (Figure 3.4b and 3.4d). A 2×3 Rm-ANOVA on the mean amplitude of the N1 component from 100-110 ms only revealed a significant main effect of condition over the temporo-parietal ($F(2,34) = 4.983$, $p<0.05$) region. Post-hoc tests using Bonferroni's correction for multiple comparisons revealed that the main effect of condition was accounted for by a significant ($p<0.05$) suppression of N1 responses when subjects prepared to either raise or lower the pitch of their voice in response to visual cues compared with control (no vocal change). This effect was robustly present in response to both upward and downward pitch-shift stimuli during the preparatory phase of vocal production (bar plots in Figures 3.4a and 3.4c).

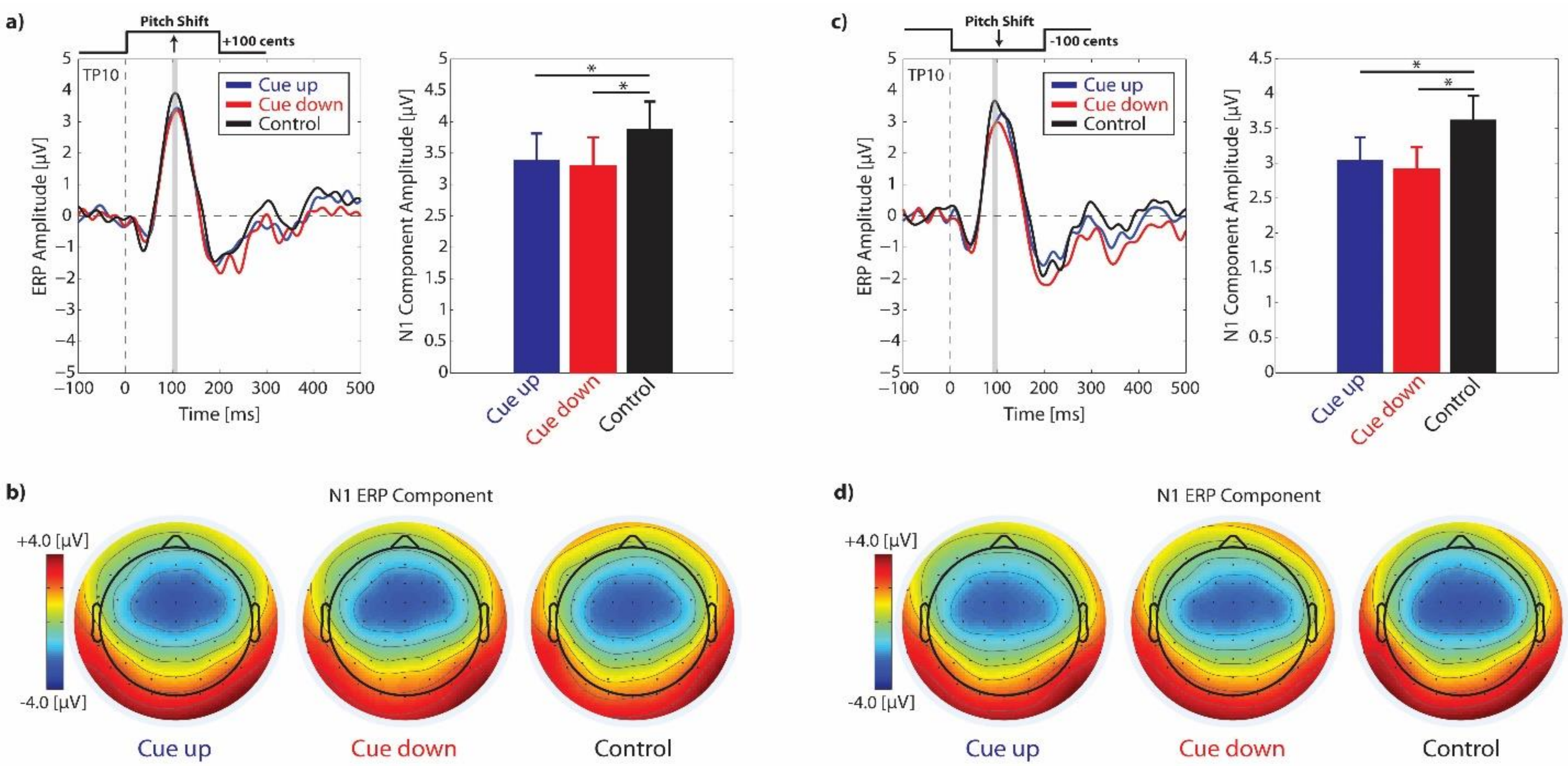

Figure 3.4 N1 ERP Response a) Profile of the overlaid grand-average N1 ERP responses to upward (+100 cents) pitch-shift stimulus for cue up, cue down and control conditions. The bar plots show the statistical comparison of the N1 component of ERPs (100-110 ms) across the three conditions. b) Topographical scalp distribution maps of the N1 component in response to upward pitch-shift stimuli for cue up, cue down and control conditions. c) Profile of the overlaid grand-average N1 ERP responses to downward (-100 cents) pitch-shift stimulus for cue up, cue down and control conditions. The bar plots show the statistical comparison of the N1 component of ERPs (40-50 ms) across the three conditions. d) Topographical scalp distribution maps of the N1 component in response to downward pitch-shift stimuli for cue up, cue down and control conditions.

3.5 P2 ERP COMPONENT

The P2 ERP responses were elicited at approximately 200 ms after the onset of pitch-shift stimuli in the auditory feedback. The profile of the grand-average ERP responses to upward and downward stimuli, overlaid across cue up, cue down and control conditions are shown in Figure 3.5. The topographical distribution maps of the P2 responses showed that this component was elicited with a positive peak of activation predominantly over the central region and inverted (negative polarity) over the temporal region (Figure 3.5b and 3.5d). A 2×3 Rm-ANOVA on the mean amplitude of the P2 component from 190-200 ms only revealed a significant main effect of condition over the central (F(2,34) = 4.886, $p<0.05$) region. Post-hoc tests using Bonferroni's correction for multiple comparisons revealed that the main effect of condition was accounted for by a significant ($p<0.05$) suppression of P2 responses only in response to upward pitch-shift stimuli for cue up compared with cue down condition (Figure 3.5a). However, in contrast with the P1 and N1 components, no such an effect was observed in response to downward stimuli for P2.

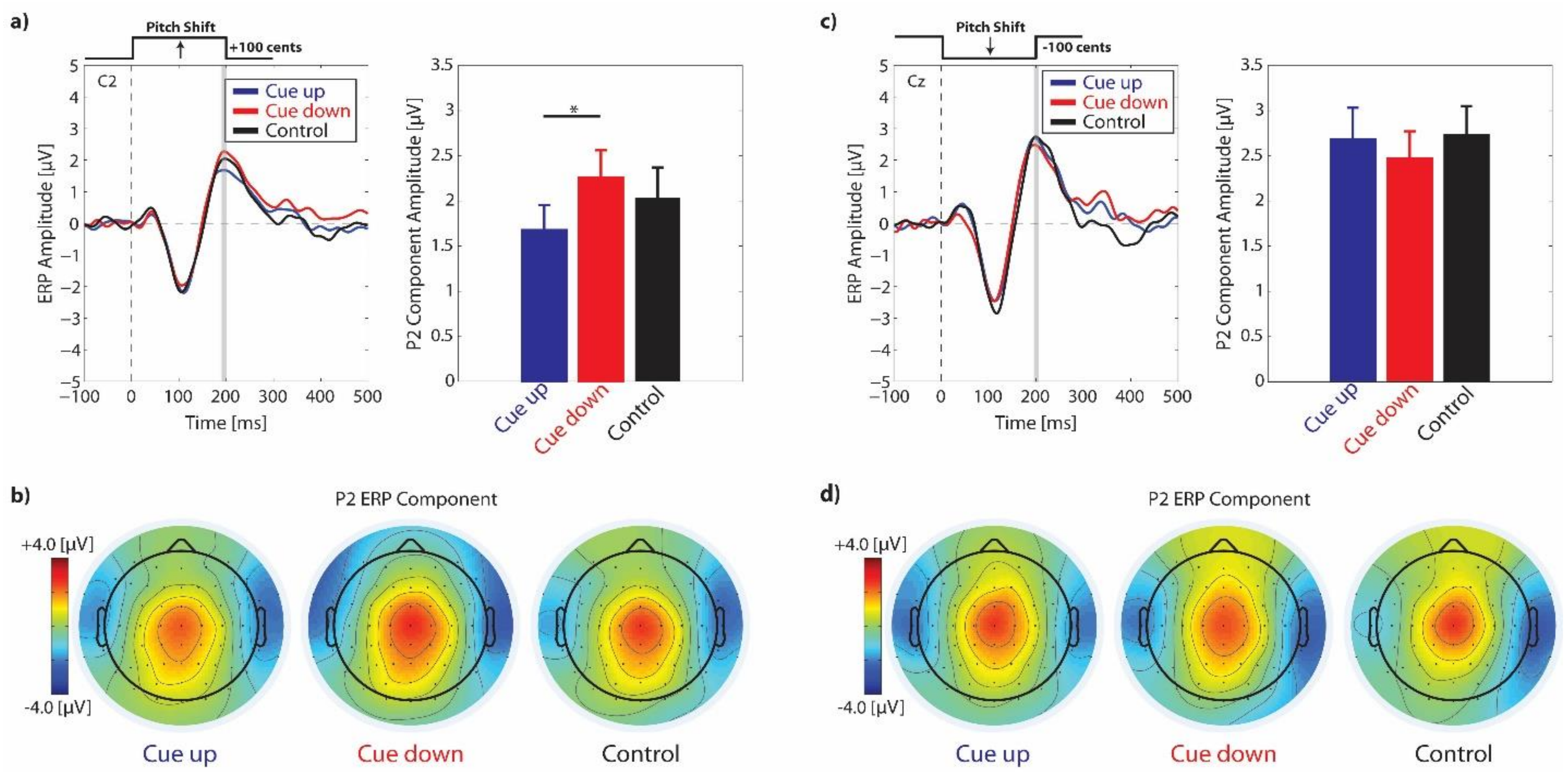

Figure 3.5 P2 ERP Response a) Profile of the overlaid grand-average P2 ERP responses to upward (+100 cents) pitch-shift stimulus for cue up, cue down and control conditions. The bar plots show the statistical comparison of the P2 component of ERPs (190-200 ms) across the three conditions. b) Topographical scalp distribution maps of the P2 component in response to upward pitch-shift stimuli for cue up, cue down and control conditions. c) Profile of the overlaid grand-average P2 ERP responses to downward (-100 cents) pitch-shift stimulus for cue up, cue down and control conditions. The bar plots show the statistical comparison of the P2 component of ERPs (40-50 ms) across the three conditions. d) Topographical scalp distribution maps of the P2 component in response to downward pitch-shift stimuli for cue up, cue down and control conditions.

CHAPTER 4

DISCUSSION

This study investigated the generation of predictive motor plans during the planning phase of speech production and the influence of these plans on voice motor control. Subjects produced repeated steady vocalizations of a vowel sound and were visually presented with a directional cue (black arrow) that instructed them to glide their pitch in the upward or downward direction, or maintain the steady vocalization (control condition) following presentation of a "go" cue. While subjects vocalized, a randomized brief (~ 200 ms) upward or downward pitch shift at 100 cents perturbed auditory feedback during the planning phase of the task. We hypothesized the internal forward model is active, and the efference copies are present, during the planning phase of vocal motor movement. We also hypothesized that neurophysiological (i.e., ERP) and behavioral response mechanisms are modulated during this planning phase.

Behavioral data revealed compensatory (opposing) vocal responses with differences in onset latency time and peak magnitude relative to the direction of the stimulus. A significant delay in onset latency of vocal response was observed for conditions that loaded a motor plan (i.e., cue up or cue down). This was observed for the upward matched cue and stimulus condition and for both the matched and mismatched cue and downward stimulus conditions. This delay may be attributed to an increased computational load (i.e., the motor plan) resulting in increased time required to process additional sensorimotor information (i.e., processing the auditory stimulus and initiating compensatory motor commands). Additionally, a significant decrease in peak vocal magnitude was observed for the matched upward stimulus and cue condition. This decrease in magnitude of compensatory response to the matching direction of stimulus and cue was consistent with findings in Ahmadi-Pajouh et al.'s (2012) limb study. Their study suggested this pattern of response was indicative of an active feedback controller during the planning phase regulating motor movements specific to the goal of the task. In our study, however, this behavioral response was only observed for the upward shift direction. No differences in peak magnitude were observed when a downward stimulus was applied. It is possible that the variability of onset latency and peak magnitude responses between upward and downward stimulus conditions may be due to differences in mechanisms of the muscle groups engaged in producing the compensatory responses. However, the framework underlying the behavioral mechanism of speech motor planning needs further exploration.

Analysis of neurophysiological responses revealed significant modulation of the P1-N1-P2 ERP components. The P1 component (also known as P50 in some literature) is believed to reflect sensory gating, or screening of redundant auditory information (Clementz et al., 1997; Korzyukov et al., 2007). P1 suppression has been repeatedly observed in paired click stimulus studies following the second presentation of an identical stimulus (Clementz et al., 1997). Interestingly, our findings revealed direction-specific modulation of the P1 ERP component evidenced by significant suppression of the P1 component when the direction of the stimulus matched the direction of the cue. We believe this modulation is evidence that, consistent with the internal forward model theory, the motor plan (efference copies) and the predicted sensory consequences (corollary discharges) for the cued vocal task were loaded during the planning phase of following the initial presentation of the directional cue. The stimulus that matched the sensory prediction was interpreted by the brain as redundant information resulting in suppression of the P1 ERP component.

Significant N1 suppression was observed for trials in which an auditory change was anticipated (i.e., subjects were cued to change their pitch), regardless of the direction. This observed suppression is consistent with other studies that suggest N1 suppression occurs when there is a match between internally generated sensory predictions and actual sensory feedback (Behroozmand et al., 2009; Ford et al., 2014; Korzyukov et al., 2012). P2 responses did not reveal as steady a pattern as the P1 and N1 responses. Rather, analysis revealed modulation of the P2 component only for the upward stimulus direction. Significant suppression of P2 was only observed when there was a match between the upward stimulus and cue as compared to the mismatched upward stimulus and cue. A similar trend was observed in conditions with for the downward stimulus direction; however, the suppression did not reach significance.

Notable differences were observed between the sensitivities of P1, N1, and P2 components in this study. The P1 component was most responsive to the vocal motor plan and appeared to reflect direction-specific neural processing of pitch evidenced by a significant level of suppression for conditions in which the direction of the cue and stimulus matched as compared to conditions with a directional mismatch. The N1 component was also sensitive to the loading of a motor plan evidenced by significant N1 suppression for conditions in which subjects were cued to change their pitch as compared to control. However, N1 modulation did not appear to be a function of congruence or incongruence between cue and stimulus direction as was observed in the P1 response. The P2 component exhibited the least sensitivity to motor planning compared to P1 and N1 responses. Interestingly, P2 suppression was observed when the direction of the stimulus was congruent with the cue only in response to upward shifts. This response was similar to behavioral data for the peak magnitude of vocal compensations possibly reflecting P2 involvement in motor rather than auditory modulation.

Our study revealed that topographical distribution maps of the P1 responses elicited a positive peak of activation predominantly over the frontal region and inverted polarity over the temporo-parietal region. N1 responses elicited a positive peak of activation predominantly over the fronto-central region also with an inverted polarity over the temporo-parietal region. P2 responses triggered a positive peak of activation predominantly over the central region with inverted polarity over the temporal region. Possible neural generators of these auditory P1-N1-P2 ERP responses have been suggested in previous fMRI and source reconstruction studies. These studies proposed generation of these components originate from the frontal, temporal, and/or parietal lobes from specific areas such as the inferior frontal gyrus, superior temporal gyrus, Heschl's gyrus, and supplementary motor area (Behroozmand et al., 2015; Korzyukov et al., 2007; Martin et al., 2008; Parkinson et al., 2012; Wang et al., 2014). In the context of our study, we propose that P1 is generated by activity from the primary auditory cortex with increased sensitivity to the frequency of the auditory stimulus. The N1 and P2 components appear to reflect higher-level sensorimotor processing possibly generated by primary and secondary auditory cortices with motor area involvement.

The outcome of this research has implications for elucidating the underlying mechanisms of clinical populations with neurological disorders. Studies indicate a disruption in feedforward and feedback mechanism in patients with Parkinson's disease (Carlsen et al., 2012; X. Chen et al., 2013; Liu et al., 2012; Mollaei et al., 2013), stuttering (Cai et al., 2012; Daliri and Max, 2015; Loucks et al., 2012), and schizophrenia (Ford et al., 2001, 2014). This disruption is suggested to contribute to symptoms such as auditory verbal hallucinations in patients with schizophrenia or dysfluencies in people who stutter. Our research may provide an objective approach to probe the integrity of the sensorimotor system and observe consequences of treatments on the function of the motor network in these clinical populations.

Further study is warranted to better understand the mechanism of behavioral responses and identify specific neuroanatomical generators of the observed ERP components. It would also be of interest examine neurophysiological and behavioral responses to stimuli at time points in the planning phase that occur closer to the "go" cue. Our study applied the stimulus at random times during a time window of 750-1000 ms before the "go" cue. The limb study by Ahmadi-Pajouh et al. (2012) found that, prior to adaptation to a specific force field, subjects produced smaller hand displacement when a perturbing pulse was applied -350 ms before reach onset as compared to when a pulse was applied at -750 ms before reach onset. Examination of this temporal aspect in the speech modality would provide additional insight to the loading of predictive plans during the planning phase of a speech task.

In closing, we propose the findings of this study provide evidence that internal sensorimotor predictions are loaded during the planning phase of a cued vocal pitch modulation task. Further, in line with the internal forward model theory, we propose that the production of efference copies of the motor commands (i.e., increase, decrease, or maintain vocal pitch following a “go” cue) give rise to corollary discharges of the expected sensory consequences of these self-produced vocalizations (i.e., a change in auditory feedback) which negates neural responses consistent with the anticipated sensory changes. This sensorimotor plan modulates neurophysiological and behavioral responses to pitch changes in voice auditory feedback as evidenced by suppression of P1-N1-P2 ERP responses for conditions requiring a motor plan, and the modulation of onset latency and peak vocal magnitude of behavioral responses during the planning phase of the vocal task.

REFERENCES

Ahmadi-Pajouh, M. A., Towhidkhah, F., & Shadmehr, R. (2012). Preparing to Reach: Selecting an Adaptive Long-Latency Feedback Controller. *Journal of Neuroscience*, *32*(28), 9537–9545. http://doi.org/10.1523/JNEUROSCI.4275-11.2012

Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, 12(1), 54. http://doi.org/10.1186/1471-2202-12-54

Behroozmand, R., Karvelis, L., Liu, H., & Larson, C. R. (2009). Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clinical Neurophysiology*, *120*(7), 1303–1312. http://doi.org/10.1016/j.clinph.2009.04.022

Behroozmand, R., Korzyukov, O., Sattler, L., & Larson, C. R. (2012). Opposing and following vocal responses to pitch-shifted auditory feedback: Evidence for different mechanisms of voice pitch control. *The Journal of the Acoustical Society of America*, *132*(4), 2468. http://doi.org/10.1121/1.4746984

Behroozmand, R., Shebek, R., Hansen, D. R., Oya, H., Robin, D. A., Howard, M. A., & Greenlee, J. D. W. (2015). Sensory–motor networks involved in speech production and motor control: An fMRI study. *NeuroImage*, 109, 418–428. http://doi.org/10.1016/j.neuroimage.2015.01.040

Boersma, P., Weenik, D. (1996). PRAAT: a system for doing phonetics by computer. Rep. Inst. Phonetic Sci. Univ. Amsterdam.

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, *103*(6), 3153–3161. http://doi.org/10.1121/1.423073

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE*, *7*(7), 1–14. http://doi.org/10.1371/journal.pone.0041830

Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2011). Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(45), 16483–16490. http://doi.org/10.1523/JNEUROSCI.3653-11.2011

Carlsen, A. N., Maslovat, D., & Franks, I. M. (2012). Preparation for voluntary movement in healthy and clinical populations: Evidence from startle. *Clinical Neurophysiology*, *123*(1), 21–33. http://doi.org/10.1016/j.clinph.2011.04.028

Chen, X., Zhu, X., Wang, E. Q., Chen, L., Li, W., Chen, Z., & Liu, H. (2013). Sensorimotor control of vocal pitch production in Parkinson's disease. *Brain Research*, *1527*, 99–107. http://doi.org/10.1016/j.brainres.2013.06.030

Chen, Z., Chen, X., Liu, P., Huang, D., & Liu, H. (2012). Effect of temporal predictability on the neural processing of self-triggered auditory stimulation during vocalization. *BMC Neuroscience*, *13*(1), 55. http://doi.org/10.1186/1471-2202-13-55

Clementz, B. A., Geyer, M. A., & Braff, D. L. (1997). P50 suppression among schizophrenia and normal comparison subjects: a methodological analysis. *Biological Psychiatry*, 41(10), 1035–44. http://doi.org/10.1016/S0006-3223(96)00208-9

Daliri, A., & Max, L. (2015). Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain and Language*, 143, 59–68. http://doi.org/10.1016/j.bandl.2015.03.002

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1), 9-21.

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological Evidence of Corollary Discharge Dysfunction in Schizophrenia. *American Journal of Psychiatry*, *158*(12), 2069. Retrieved from https://login.pallas2.tcl.sc.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=5667355&site=ehost-live

Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2014). Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone. *Schizophrenia Bulletin*, 40(4), 804–812. http://doi.org/10.1093/schbul/sbt072

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, 39(5), 350–365. http://doi.org/10.1016/j.jcomdis.2006.06.013

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96(3), 280–301. http://doi.org/10.1016/j.bandl.2005.06.001

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, 42(2), 180–190. http://doi.org/10.1111/j.1469-8986.2005.00272.x

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*, *17*(13), 1375–9. http://doi.org/10.1097/01.wnr.0000233102.43526.e9

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69(3), 407–422. http://doi.org/10.1016/j.neuron.2011.01.019

Houde, J. F., & Chang, E. F. (2015). The cortical computations underlying feedback control in vocal production. *Current Opinion in Neurobiology*, *33*, 174–181. http://doi.org/10.1016/j.conb.2015.04.006

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5, 82. http://doi.org/10.3389/fnhum.2011.00082

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience*, 14(8), 1125–1138. http://doi.org/10.1162/089892902760807140

Korzyukov, O., Pflieger, M. E., Wagner, M., Bowyer, S. M., Rosburg, T., Sundaresan, K., … Boutros, N. N. (2007). Generators of the intracranial P50 response in auditory sensory gating. *NeuroImage*, 35(2), 814–826. http://doi.org/10.1016/j.neuroimage.2006.12.011

Korzyukov, O., Sattler, L., Behroozmand, R., & Larson, C. R. (2012). Neuronal mechanisms of voice control are affected by implicit expectancy of externally triggered perturbations in auditory feedback. *PLoS ONE*, *7*(7). http://doi.org/10.1371/journal.pone.0041216

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. The Journal of *Neuroscience: The Official Journal of the Society for Neuroscience*, 32(27), 9351–9358. http://doi.org/10.1523/JNEUROSCI.0404-12.2012

Liu, H., Wang, E. Q., Metman, L. V., & Larson, C. R. (2012). Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson's disease. *PLoS ONE, 7*(3). http://doi.org/10.1371/journal.pone.0033629

Loucks, T., Chon, H., & Han, W. (2012). Audiovocal integration in adults who stutter. *International Journal of Language and Communication Disorders*, *47*(4), 451–456. http://doi.org/10.1111/j.1460-6984.2011.00111.x

Martin, B. A., Tremblay, K. L., & Korczak, P. (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear and Hearing*, 29(3), 285–313. http://doi.org/10.1097/AUD.0b013e3181662c0e

Max, L., Daniels, J., Curet, K., & Cronin, K. (2008). Modulation of auditory and somatosensory processing during the planning of speech movements. Paper presented at the Proceedings of the 8th international seminar on speech production, Strasbourg, France (pp. 41–44).

Mollaei, F., Shiller, D. M., & Gracco, V. L. (2013). Sensorimotor adaptation of speech in Parkinson's disease. *Movement Disorders, 28*(12), 1668–1674. http://doi.org/10.1002/mds.25588

Parkinson, A. L., Flagmeier, S. G., Manes, J. L., Larson, C. R., Rogers, B., & Robin, D. A. (2012). Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage*, 61(1), 314–322. http://doi.org/10.1016/j.neuroimage.2012.02.068

Scheerer, N. E., & Jones, J. A. (2014). The predictability of frequency-altered auditory feedback changes the weighting of feedback and feedforward input for speech motor control. *European Journal of Neuroscience*, *40*(12), 3793–3806. http://doi.org/10.1111/ejn.12734

Scheerer, N. E., Behich, J., Liu, H., & Jones, J. A. (2013). ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience*, 240, 176–185. http://doi.org/10.1016/j.neuroscience.2013.02.054

Wang, J., Mathalon, D. H., Roach, B. J., Reilly, J., Keedy, S. K., Sweeney, J. A., & Ford, J. M. (2014). Action planning and predictive coding when speaking. *NeuroImage*, *91*, 91–98. http://doi.org/10.1016/j.neuroimage.2014.01.003

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729–R732. http://doi.org/10.1016/S0960-9822(01)00432-8

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An Internal Model for Sensorimotor Integration. *Science*, *269*(5232), 1880–1882. Retrieved from https://login.pallas2.tcl.sc.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=11016183&site=ehost-live