THE NEURAL SUBSTRATES OF SENSORY-MOTOR CONTROL OF THE VOICE: A NON-INVASIVE IMAGING APPROACH

by

Jordan L. Manes

THESIS
Presented to the Faculty of the
Honors College
The University of Texas at San Antonio
In Partial Fulfillment
Of the Requirements
For the Degree of

BACHELOR OF ARTS IN PSYCHOLOGY
WITH HIGHEST HONORS IN THE HONORS COLLEGE

THE UNIVERSITY OF TEXAS AT SAN ANTONIO
College of Liberal and Fine Arts
Department of Psychology
May 2011

THE NEURAL SUBSTRATES OF SENSOR-MOTOR CONTROL OF THE VOICE: A NON-INVASIVE IMAGING APPROACH

PREPARED BY:

_____
Jordan L. Manes

APPROVED BY:

_____
Donald A. Robin, Ph.D., Thesis Advisor

_____
Charles Larson, Ph.D., Thesis Reader

_____
Ann Eisenberg, Ph.D., Thesis Reader

Accepted: _____
Richard Diem, Ph.D., Dean of the Honors College

Received by the Honors College:

_____

ACKNOWLEDGEMENTS

My deepest gratitude goes out to all those who have contributed to this project. I would like to thank Donald A. Robin for his endless support and encouragement, along with all the members of the Human Performance Lab at the UT Health Science Center at San Antonio. I would also like to extend my thanks to Chuck Larson and Ann Eisenberg for their assistance in pulling this together. Working with all of you has been a wonderful privilege; and I am ever grateful for your guidance.

May 2011

ABSTRACT

THE NEURAL SUBSTRATES OF SENSOR-MOTOR CONTROL OF THE VOICE: A NON-INVASIVE IMAGING APPROACH

Jordan Manes, B.A.
The University of Texas at San Antonio, 2011

Supervising Professor: Donald A. Robin, Ph.D.

Many communication disorders arise from disruptions in the neurological ability to control the voice. Deficits in sensorimotor integration can contribute to these disorders through disruption of audio-vocal reflexes; however the neural mechanisms underlying audio-vocal reflexes are still unknown. While our previous research has successfully used the pitch-shift paradigm to characterize these mechanisms in healthy adults, it has failed to account for differing vocal responses to the pitch-shift task (i.e., following vs. compensatory responses).This study first reviews our previous research using the pitch-shift paradigm in conjunction with fMRI and TMS. It then expands these findings by comparing fMRI activations patterns across response types and using a non-speech control site to validate previous TMS findings. During fMRI, subjects were asked to sustain a steady phonation for 5 second intervals, during which feedback of vocal pitch was shifted up/down by 100 cents. While subjects performed this task, fMRI images were acquired for Rest, Vocalize (No shift), and Vocalize (Shift) conditions. In the present study, we classified each pitch-shift vocalization trial as a compensatory, following, or mixed response and performed subsequent contrast analyses comparing following and compensatory responses. In compensatory trials, stronger activation was seen in bilateral insula,

left STG, and right IFG. Following trials were associated with stronger activation in the right inferior parietal cortex, right superior occipital gyrus, right precuneus, posterior cingulate, middle cingulate, bilateral superior frontal gyrus, left middle frontal gyrus, and middle orbital gyrus. In our previous work, we also attempted to quantify the timing of M1 involvement in the pitch-shift reflex. Subjects again performed the same task while undergoing TMS of the primary motor cortex at delay times of 0, 10, 20, 30, 40, 50, and 60 ms after onset of the pitch-shift stimulus. Although we were able to observe an interaction effect at 30ms after stimulus onset, there was no TMS control site to assure this effect was site-specific and not a main effect of TMS. We thus used TMS in this study to stimulate a non-speech control site (vertex) during vocalization. As predicted, stimulation of vertex led to no difference in voice F0 during vocalization, validating our previous work on the timing of M1 involvement in the pitch-shift reflex.

TABLE OF CONTENTS

ABSTRACT ..... **V**

LIST OF FIGURES ..... **VIII**

CHAPTER 1: INTRODUCTION **1**

CHAPTER 2: THESIS STATEMENT **10**

2.1 HYPOTHESIS 1 10
2.2 HYPOTHESIS 2 10

CHAPTER 3: METHODS AND APPROACH **12**

3.1. SUBJECTS 12
3.2. PITCH SHIFT 12
3.3. PROCEDURE: FMRI 13
3.4. PROCEDURE: TMS 16
3.5. LIMITATIONS OF METHODS AND APPROACH 18

CHAPTER 4: RESULTS AND DISCUSSION **19**

4.1. SUMMARY OF PREVIOUS FMRI STUDY 19
4.2. RESULTS OF CURRENT FMRI STUDY 20
4.3. SUMMARY OF PREVIOUS TMS STUDY 23
4.4. RESULTS OF CURENT TMS STUDY 24
4.6. DISCUSSION 25

CHAPTER 5: CONCLUSIONS **32**

CHAPTER 6: IMPLICATIONS OF RESEARCH **33**

REFERENCES **35**

LIST OF FIGURES

Figure 1.1 Voice F0 response to pitch-shifted feedback 4

Figure 3.1 Timeline of a single fMRI trial 15

Figure 4.1 BOLD activations for Vocalize vs. Rest conditions 19

Figure 4.2 BOLD activations for Shift vs. No Shift conditions 20

Figure 4.3 Percentages of compensatory, following, and mixed response trials 21

Figure 4.4 BOLD contrast of Compensatory vs. Following responses 22

Figure 4.5 BOLD contrast of Following vs. Compensatory responses 22

Figure 4.6 F0 response magnitude to pitch-shifted feedback during TMS 23

Figure 4.7 Effects of TMS stimulation of M1 on F0 during vocalization 25

Figure 4.8 Effects of TMS stimulation of vertex on F0 during vocalization 25

CHAPTER 1: INTRODUCTION

The human voice has unique qualities and is used for expression of speech and emotions as well as for many professional endeavors, such as singing or news casting. However, it is estimated that 7.5 million people in the United States alone have difficulty using their voices (National Institute on Deafness and Other Communication Disorders). Voice disorders can have profound impacts on an individual's ability to communicate and may substantially decrease the quality of life for that person. In several of these disorders, impairments in voice control can be attributed to neurological diseases (e.g., Parkinson's disease, spasmodic dysphonia,) that lead to deficits in the control of voice fundamental frequency (F0) and intensity (Duffy, 2005). Motor control and the anomalies associated with sensory motor integration are critical components of these deficits. Recent studies in both voice and general motor control call for a critical evaluation of sensory contributions to motor control, particularly in the area of sensory motor integration (SMI) (Poizner et al., 2000; Tunik, Poizner, Adamovich, Levin, & Feldman, 2004). While current research supports a model of sensorimotor integration for voice control, the neural mechanisms by which sensory feedback contribute to the regulation of voice are still unknown (Hammer & Barlow, 2010). The sections below provide a review of sensorimotor integration and the pitch-shift reflex as well as literature on functional imaging and TMS stimulation of voice-oriented regions in the brain.

1.1. Sensorimotor-integration in motor systems including voice

1.1.1. SMI Background

It is well known that sensory and perceptual factors play a substantial role in the control of motor output. Being deprived of visual input, for instance, leads to deficits in our sense of limb position and overall proprioception (Gosselin-Kessiby, Kalaska, & Messier, 2009; Wolpert,

Ghahramani, & Jordan, 1995). Although voice and limb control are distinctly separate forms of motor output, the production of human speech must be recognized as a motor action that similarly relies on sensory feedback and feed-forward mechanisms of error detection (i.e., corollary discharge) (Gracco, 1991). Here we review some of the literature on sensorimotor integration (SMI) as it applies to general motor control as well as literature on SMI that is specific to the audio-vocal reflex. Below is a discussion of critical issues in SMI including efference copy in relation to vocal production. Such a mechanism is common in models of motor control and computational models of speech (e.g.,(Guenther, Ghosh, & Tourville, 2006)).

1.1.2. Efference copy

The ability to predict the outcome of an executed motor task is crucial to the control of motor output. The proposed feed-forward mechanism for error detection is known as efference copy (or corollary discharge). Efference copy provides information on performed actions that may then be used to attenuate sensory consequences thereof (Sommer & Wurtz, 2002, 2008). Theoretical neurobiology has formalized this observation within a predictive coding framework, conceptualizing the brain as a multi-hierarchical, Bayesian inference machine. This Bayesian approach specifies that predictions about sensory data and their causes be consistently formed based on prior information, such as corollary discharges from the motor system (Kilner, Friston, & Frith, 2007). While a valid prediction will devalue the information content of the incoming data, a large prediction error is thought to trigger the allocation of attentional resources, behavioural adaptation and readjustment of predictions (K. Friston, 2002; K. J. Friston & Price, 2001). This predictive framework is crucial to understanding the role of sensory feedback in motor control and its breakdown in disordered states such as Parkinson's disease.

1.2. Pitch-shift paradigm

Audio-vocal reflexes work to regulate various components of vocalization (e.g., loudness, pitch, etc.). Proper control of vocal pitch is crucial to communication in non-tonal languages, functioning as a means of emotional expression (Banziger & Scherer, 2005), indicating a question (W. E. Cooper, Eady, & Mueller, 1985; Eady & Cooper, 1986; Xu & Xu, 2005), and production of stressed syllables. When subjects are presented with unexpected perturbations in voice pitch feedback, studies have shown that they produce compensatory changes in voice F0 – a response referred to as the pitch-shift reflex. Several studies have demonstrated the importance of this reflex in the control and stabilization in voice F0 during sustained vowels (Bauer & Larson, 2003; Chen, Liu, Xu, & Larson, 2007; Hain et al., 2000; Sivasankar, Bauer, Babu, & Larson, 2005) as well during speech (Chen, et al., 2007; Xu, Larson, Bauer, & Hain, 2004), and the production of nonsense syllables (Donath, Natke, & Kalveram, 2002; Natke, Donath, & Kalveram, 2003; Natke & Kalveram, 2001).

The pitch-shift paradigm has been used extensively to investigate the role of the audio-vocal reflex in controlling voice F0 (e.g., (Hain, et al., 2000)). Within this paradigm, subjects are asked to vocalize while their voice is fed back to them through a set of headphones. As the subjects vocalize, pitch of voice auditory feedback is perturbed either upward or downward by a certain degree, typically ±100 cents (100 cents = 1 semitone). Presentation of the downward pitch-shift stimulus is then followed by an automatic compensatory response characterized by an increase in the subject's voice F0. Likewise, upward pitch-shifted feedback is accompanied by a reflexive decrease in F0. Perturbation paradigms have also been successfully applied to intensity by shifting the loudness of auditory feedback (e.g., (Bauer, Mittal, Larson, & Hain, 2006); however the control of voice loudness through sensory feedback will not be encompassed in this study. Figure 1.1 shows responses to upward (red) and downward (blue) pitch shifted stimuli.

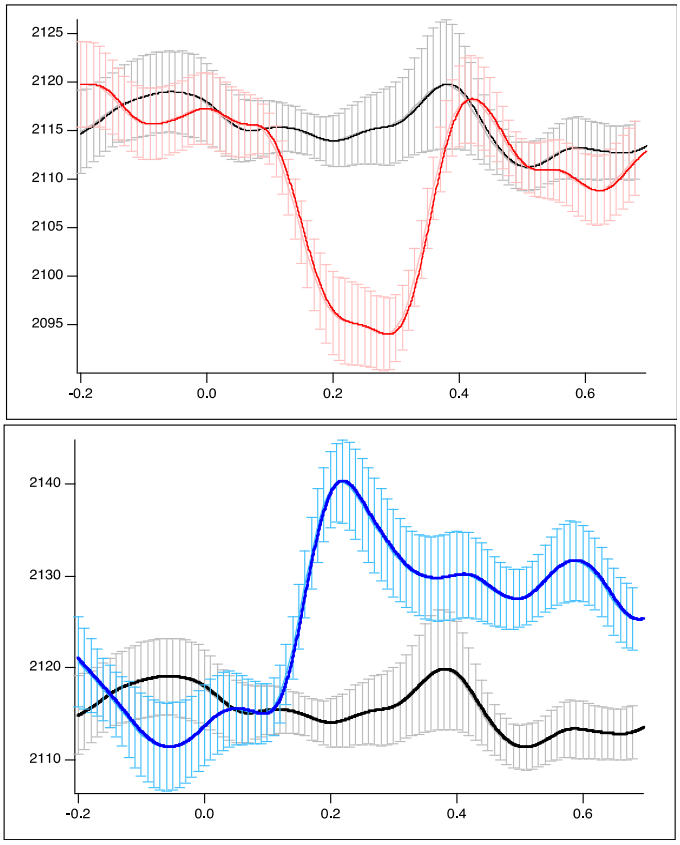


Figure 1.1 Voice F0 response to upward pitch shift (above) and downward pitch-shift (below) stimuli. Black lines represent baseline vocalization without pitch-shift. Responses to upward and downward shifted stimuli are indicated by red and blue lines, respectively. The x-axis corresponds to time (s). The y-axis corresponds to voice F0 (Hz).

The magnitude and latencies of the pitch-shift response are dependent on the magnitude and direction of the shift (Bauer, et al., 2006), the duration of the shift (Burnett, Freedland, Larson, &

Hain, 1998), and the nature of the task being performed (Chen, et al., 2007). Chen and colleagues examined task-dependent effects of pitch perturbation by comparing F0 responses of speech versus sustained vowel tasks. They found that subjects produced larger and faster responses to the pitch-shift stimulus when performing a speech task compared to a sustained vowel task. The authors proposed that this exaggerated effect during speech tasks was due to the greater importance of pitch control in communicating effectively to a listener as opposed to sustaining vowels which do not demand effective communication (Chen, et al., 2007). Duration of pitch shift also appears to be an important factor, as stimuli exceeding the time frame of 300ms begin to elicit a voluntary response as opposed to reflexive audio-vocal control (Burnett, et al., 1998; Hain, et al., 2000). Given that the shifted auditory feedback in voice production causes the perceived sensory consequences to diverge from expected outcomes, the study of this reflex using neuroimaging will provide insight in efference copy and shed light on the mechanisms of action of the phenomenon.

In some instances, voice F0 changes in the same direction as the pitch-shift stimulus (i.e., an upward change in F0 in response to an upward pitch-shift), an effect referred to as “following”. While this type of response is relatively rare, it appears to occur more frequently with greater perturbation of voice F0 (Burnett, et al., 1998). A possible interpretation of this effect is that when voice feedback is distorted beyond a particular threshold, it becomes recognized as an “alien voice” as opposed to a self-generated sound. Without perceived change in self-produced F0, such responses would not reflect the process of SMI in the audio-vocal system. As such, further investigation comparing following versus compensatory responses might be useful in distinguishing one as indicative of SMI and the other as an atypical response to pitch perturbations.

1.3. Functional Imaging

1.3.1. Blood Oxygen Level Dependent (BOLD) Response

While traditional MR imaging utilizes the paramagnetic properties of various tissues to create anatomical images, functional MRI utilizes the diamagnetic properties of oxyhemoglobin and paramagnetic properties of deoxyhemoglobin to measure changes in the hemodynamic response and image activity in the brain. When an area of the brain is active, a proportional increase in the supply of oxygenated hemoglobin is delivered to that region. Since diamagnetic substances have a higher T2* than paramagnetic substance, regional increases in oxyhemoglobin result in a stronger MR signal, thereby allowing us to image activity in the brain. (McRobbie, 2003).

Functional imaging has been used extensively to examine an array of voice related tasks, including speech, singing, and reading; however, very few have examined phonation as a distinct entity apart from the general speech system. Brown and colleagues (2009; 2008) have recently conducted functional imaging studies examining the cortical areas involved in phonation. As a result of their work, they have proposed a somatotopic region of the primary motor cortex that corresponds to the larynx, an area they refer to as the laryngeal motor cortex (LMC). The proposed LMC is located in the same general area that has been implicated in the motor control of expiratory muscles (Loucks, Poletto, Simonyan, Reynolds, & Ludlow, 2007). Since breathing and laryngeal control are fundamental to phonation and the control of voice pitch, this region of the motor strip is likely to be to be key in regulating those components of speech.

In addition, Brown (2009) also sought to differentiate the various components of speech by incorporating speech, orofacial, tongue, and phonation tasks. The results supported their previous research implicating the laryngeal motor cortex (LMC) as the main cortical area involved in

phonation. In this work, the Rolandic operculum (RO) was active during the control of tongue movement and speech articulation. These results were further combined with a quantitative meta-analysis of syllable singing and reading tasks to develop a neural model of vocalization. The proposed network specific to phonation includes the primary and premotor cortices, supplemental motor area, posterior temporal gyrus, and cerebellum.

1.4. TMS

The TMS technique requires that the pulse be precisely delivered to the location of the putative neurons involved in the reflex. To accomplish this objective, the Tailarach coordinates, as derived from the fMRI analysis, are fed to a computer that controls a robotic arm that holds the magnetic coil of the TMS (irTMS) unit. The irTMS coil is then guided by the computer controlled robot to a position over the targeted area in an orientation that most effectively activates the neurons in that region. The technique of TMS has been used to provide important information on the function of different cortical areas related to reflexes (A. C. Cooper, Humphreys, Hulleman, Praamstra, & Georgeson, 2004; Fisher, Sharott, Kuhn, & Brown, 2004; Ito, Kimura, & Gomi, 2005; Petersen, Pyndt, & Nielsen, 2003) and voluntary motor control mechanisms (Tzelepi, Yang, & Kapoula, 2005) including vocalization (Lo & Fook-Chong, 2004; R. M. W. Rodel et al., 2004; Terao et al., 2001) and speech (Watkins & Paus, 2004). TMS has also been used to map cortical regions involved in speech production. Some studies (Jennum, Friberg, Fuglsangfrederiksen, & Dam, 1994; Michelucci et al., 1994; Pascualleone, Gates, & Dhuna, 1991) used a circular TMS coil that is not focal, and they could not make strong claims about localization of TMS effects. Epstein et al., (1996; 1999) localized the site of TMS induced disruption to the primary motor cortex. Stewart et al. (2001) applied TMS on two regions of the frontal lobe: a posterior site located lateral to the hand region of the primary motor cortex (MC)

and a second site located 5.5 cm anterior to the first one in the middle frontal gyrus. They found that TMS of the posterior site (presumably the mouth MC) caused speech disruption that was associated with activation of the mentalis muscle, whereas TMS of the anterior region produced a "nonmotoric" speech disruption. Importantly, our methods (irTMS) allow for precise localization of the TMS pulse to the target region using the methods described below.

Purpose

Our group has used fMRI during the pitch-shift paradigm to study the neural systems involved in response to pitch perturbation, the findings of which are summarized in the results section. While quite informative, those results are based on a combination of compensatory and following responses to pitch-shifted feedback. Because these response classes likely result in different neural system responses, it is important to differentiate compensatory responses, which are indicative of SMI in the correction of an error signal, and following responses, which indicate that correction mechanism has gone awry. Thus, the first experiment of the current study was designed to reanalyze the fMRI data that had been previously collected to identify regions of the brain involved in compensatory versus following responses to a pitch shifted stimulus.

A separate aim of our work seeks to determine the time sequence of cortical activations during the pitch-shift reflex. We have previously collected data on the timing of cortical activations through the use of transcranial magnetic stimulation (TMS), and in particular the primary motor region of the left frontal cortex. In that study we applied TMS to the laryngeal motor cortex (LMC) at delay times of 0, 10, 20, 30, 40, 50, and 60 ms after onset of the pitch-shift stimulus and found that an interaction between TMS and pitch-shift responses occurred at 30-40 ms delay time (more details are provided below). However, our initial data sets do not contain a control stimulation site that is critical in showing that TMS to regions of interest are

not occurring at any site in the brain. Therefore, the second aim of this study was to collect data on TMS to regions outside of the speech/voice network to ensure that targeted interactions were site specific.

CHAPTER 2: THESIS STATEMENT

This study is designed to delineate the neural mechanisms that are involved in the sensory control of the human voice. Using functional magnetic resonance imaging (fMRI), it will provide information on the regions involved in the sensory control of the voice; and stimulation with transcranial magnetic stimulation (TMS), which stimulates populations of neurons in a specific cortical region, will provide information about processing time between regions of the voice control network. By expanding on our previous research, we hope to further characterize the underlying neural mechanisms of the pitch-shift reflex in terms of differing behavioral responses to the task. In addition, we aim to determine whether the effects of TMS stimulation of LMC shown in our previous work are site specific or a generalized effect of TMS.

2.1 Hypothesis 1

We believe that these different classes of responses to the pitch shift stimulus reflect different neural processes – one of which reflects accurate error detection and correction within the sensorimotor integration network, and the other which indicates that this mechanism has gone awry. It is anticipated that different patterns of BOLD activation will emerge when comparing activity associated with following versus compensatory responses to the pitch shift task.

2.2 Hypothesis 2

While we have previously reported an interaction between vocal changes associated with TMS stimulation and voice response to pitch-shifted stimuli, we did not have a control stimulation site to ensure that our findings were site specific. We also have shown that F0 increases with stimulation of LMC during non-shifted vocalization. We hypothesize that

stimulation at the vertex, a neutral site that is never activated during speech tasks, will not result in any change in F0 compared to TMS stimulation of LMC where we see increases in F0 associated with TMS.

CHAPTER 3: METHODS AND APPROACH

3.1. Subjects

Ten healthy subjects were recruited locally from the University of Texas Health Science Center at San Antonio (UTHSCSA) and the University of Texas at San Antonio (UTSA). To qualify for participation, all subjects had to be right-handed English speakers with no history of language, cognitive or behavioral disorders. Additional criteria included less than five years of musical experience and normal hearing. All participants were made aware of the potential risks and benefits of participation, after which informed consent was be obtained. Subjects' hearing thresholds for pure tones were screened at 125, 250, 500, 1000, and 2000 Hz at 30 dB. Compensation consisted of $200 for the completion of the study.

3.2. Pitch-shift task

Pitch Shifting. We used experimental and analysis methods similar to those used by Chen et al., (2007). MIDI (musical instrument digital interface) software (Max/MSP, Cycling 74) on a Mac laptop, connected to a MOTU Ultralite USB Audio Interface controlled an Eventide Eclipse Harmonizer to deliver pitch- or loudness-shifted voice feedback to the subject in near real time. The software also generated TTL pulses indicating time of upward and downward pitch-shifted feedback. The pitch-shift stimulus duration was limited to 200 ms to prevent the occurrence of voluntary reactions. The pitch-shift magnitude was ± 100 cents. Subjects were told that their voice may occasionally sound unusual, but they should ignore it and just attempt to complete the task as instructed.

3.3. Magnetic Resonance Imaging

3.3.1. Image acquisition

Magnetic resonance images were acquired using a Siemens TIM TRIO 3-Tesla MRI scanner at the Research Imaging Institute of the University of Texas Health Science Center at San Antonio. T1-weighted anatomical imaging was performed using a high-resolution (isotropic 800 μm voxels) retrospective motion-corrected protocol (Kochunov et al., 2005). Six full-resolution volumes were acquired via a T1-weighted, 3-D TurboFlash sequence with an adiabatic inversion contrast pulse. The scan parameters were as follows: TE/TR/TI = 3.04/ 2100/ 785 ms, flip angle =13. The anatomic images were used to register the brain of each subject to their functional data (fMRI). Gradient-echo pulse sequences were used to acquire T2* weighted functional images. Although traditional block or event-related scanning would have been advantageous in imaging the hemodynamic response of the brain while the audio-vocal reflex is active, the presence of noise from the scanner presented methodological confounds in obtaining behavioral data of vocal responses. These include the interference of scanner noise with the collection of voice recordings and the presentation of the pitch-shift stimulus and auditory feedback. In consideration of these issues, functional images were obtained using sparse sampling – a method that captures the BOLD signal immediately after a task has been completed.

3.3.2. Procedure: fMRI

During testing, the subjects lay in the scanner and viewed a monitor screen reflected off a mirror directly above the eyes. With microphone and earphones in place, subjects were instructed to keep their jaw, lips and tongue motionless during the testing and to keep their jaw slightly open. They were also instructed that on the "vocalize" trials, they should quickly inhale and then sustain an

"ah" sound at a constant pitch and intensity until the "vocalize" instruction went off. Each trial began with an instruction on the monitor screen to either "vocalize" or to "rest" that remained illuminated for five seconds. On one half of the "vocalize" trials, three to five pitch-shifts were presented to the subject (randomized between upward and downward trials). The other "vocalize" periods did not have any stimuli presented and served as the active control vocalization. The test and control trials were randomized. After each vocalization, there was a 250 ms delay followed by a three-second scan, during which a full brain volume was obtained. There was then a four second delay following the scan to allow task-related cerebral blood flow to return to baseline before the next trial. For rest trials there was no vocalization but the timing of scanning was unchanged from vocal trials. The five second delay from vocalization onset to the beginning of scanning was used to allow the hemodynamic (BOLD) response to reach peak levels estimated to be between 4 and 7 seconds after the onset of vocalization (Tourville, Reilly, & Guenther, 2008). Auditory feedback was turned off during scanning to prevent transmission of scanner noise over the headphones. There were three blocks of 65 trials. Each trial lasted 12 seconds, and each block lasted 14.4 minutes. Following each block there was a two-minute rest period.

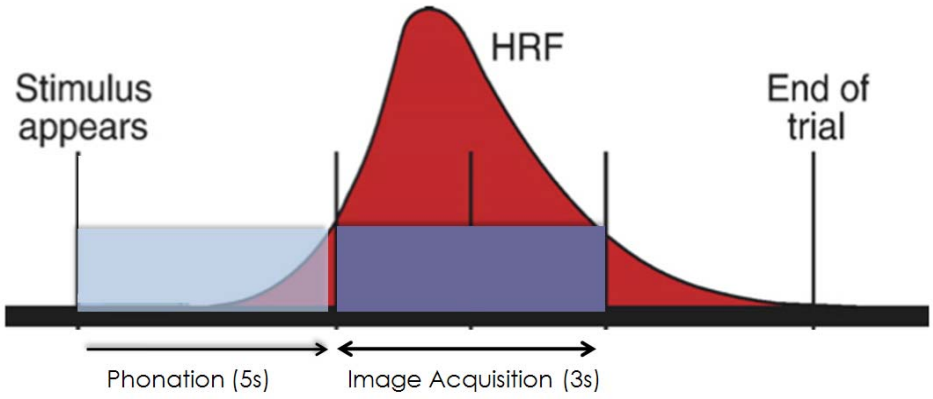


Figure 3.1. Timeline of single fMRI trial. Subjects phonate for 5 seconds after presentation of the pitch-shift stimulus. Image acquisition begins 0.25 seconds post-vocalization and imaging continues for 3 seconds. The HRF curve (red) refers to the hemodynamic response (BOLD signal).

3.3.3. Analysis of fMRI data

We performed fMRI contrast analyses using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/), subtracting the rest condition from each experimental condition and contrasting each experimental condition. Contrasts previously studied were: (vocalization no shift - rest), (vocalization pitch shift-rest), and (vocalization pitch shift – no pitch shift). As noted above, there were five pitch shifts for every vocalization in the scanner. We divided the trials into those that were mostly compensatory (4/5) or mostly following (4/5). We then analyzed the BOLD data by performing two contrasts which were (1) compensatory greater than following and (2) following greater than compensatory.

Anatomical MRI, as described above for the T1 weighted image, was registered with the fMRI data. Functional activation studies were acquired using a gradient-echo planar imaging (EPI) pulse sequence with the following parameters: TR = 12 s, TR delay = 3 s, TE = 30 ms, and a flip angle of 90 degrees. We collected 43 slices with no gap, and voxel size of 1.72 x 1.72 x 3 mm. Data were analyzed with respect to changes due to task-related differences (e.g., speech versus rest). fMRI

data were analyzed with SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). To combat movement artifact, each BOLD image in a time series underwent a 3D co-registration (6 parameter rigid body) to the middle data point in the time series (Jenkinson, Bannister, Brady, & Smith, 2002). Data were smoothed with a nonlinear algorithm designed to preserve image structure by only smoothing over voxels thought to be of the same tissue type (5 mm kernel) (Bullmore et al., 1996). For each intra-cranial voxel, least-squares coefficients were generated independently reflecting conditions in each task (Woolrich, Ripley, Brady, & Smith, 2001).

Contrast analyses of fMRI data. Within-subject statistical images were created by contrasting the vocalization conditions above with rest. For group analysis, a "minimal deformation" common space brain was defined, approximating the average size, shape and orientation of the participants (Kochunov, et al., 2005). Based on parameters created from high resolution anatomical images, statistical images created for each subject were normalized into this common space (12 parameter linear model). The common space brain was then normalized into a standard stereotactic space (Evans, 1993). An initial analysis was performed using a fixed effects model, which does not account for variability between subjects. Higher-level multi-subject analysis then utilized a mixed effects model where subject is treated as a random factor (Woolrich, et al., 2001) and provided z-images reflecting contrast specific activation patterns for each group. These group maps were thresholded at $p < .001$ based on magnitude and extent of activation (Forman et al., 1995; Poline, Worsley, Evans, & Friston, 1997), a process that statistically controls for multiple comparisons while acknowledging the improbability of cluster of contiguous "activated" voxels.

3.4. Transcranial magnetic stimulation

3.4.1. Targeting plans

irTMS was applied as described in Lancaster et al., (2004). Prior to the study day, the following procedures were performed on the MRI images. Target-site determination was based upon

co-registered anatomical (aMRI) and functional (fMRI) brain images. Coil pose was planned based on the cortical-column aiming theory (P. T. Fox et al., 2004), an orientation scheme which assumes that the greatest stimulation efficiency occurs when the TMS coil's induced E-field (y-axis) is parallel to cortical columns. The E-field for the Cadwell figure-eight coil was calculated using previously published methods (Epstein, Fernandez-Beer, Weissman, & Matsuura, 1991) and its calibration verified. The magnitude and direction of the E-field at a planned site was easily found from this model. Targeting plans were made for every subject for M1-larynx.

3.4.2. Procedure: TMS

In preparation for TMS, subjects lay on their backs on a comfortable examination bench. An adjustable head strap was then used to stabilize the subject's head during testing. Etymotic insert earphones were then inserted into the ear canals, and the microphone placed 1 inch in front of the mouth.

Once the subjects were positioned and the head immobilized, the head was registered to the MRI-based head model by first digitizing the scalp surface using a mechanical 3D digitizer (Microscribe 3DLX, Immersion Corp., San Jose, CA). The digitizer-derived scalp model was then registered to the MRI-derived scalp model using the head-and-hat algorithm (Pelizzari, Chen, Spelbring, Weichselbaum, & Chen, 1989). The robot was registered to the head using the mechanical digitizer to locate landmarks on the distal end of the arm, the location of these landmarks having been previously established in the robot's native coordinate system. In previous study TMS was then applied at three different intensities, 90%, 100% and 110% of motor threshold, to better characterize the graded intensity response for both perceptual and acoustic measures and to vary the number of neurons stimulated.

For the current study, we used data from our previous work in which subjects vocalized without pitch-shift and received TMS stimulation at an intensity of 110% motor threshold and at a rate of1 Hz. We then had 3 subjects return and used the exact same protocol to stimulate the vertex of the brain.

3.5. Limitations of Methods and Approach

There are a few methodological concerns that could arise from this study. First, since the need for clear stimulus presentation and voice recordings necessitates the use of sparse-sampling in acquiring fMRI data, the limitations of this approach must be addressed. The primary concern with the sparse sampling method is that the signal to noise ratio is greatly reduced in comparison to continuous scanning. This method is therefore accompanied by lower statistical power, making it more difficult for activations to meet the threshold for statistical significance. This is particularly important for the current study because separating trials by compensatory versus following further reduces the power.

CHAPTER 4: RESULTS AND DISCUSSION

4.1. Summary of Previous fMRI Study

fMRI data from the Rest condition were subtracted from activations in Vocalize conditions as shown in Figure 4.1 below. The vocalize conditions corresponded to activations in the pre-motor (PMC) and primary motor (M1) regions as well as the supplemental motor area (SMA), superior temporal gyrus (STG), inferior frontal gyrus (IFG), and bilateral somatosensory areas. These data align well with previous studies examining non-shifted speech (Brown, et al., 2009; Tourville, et al., 2008) and phonation (Brown, et al., 2008).

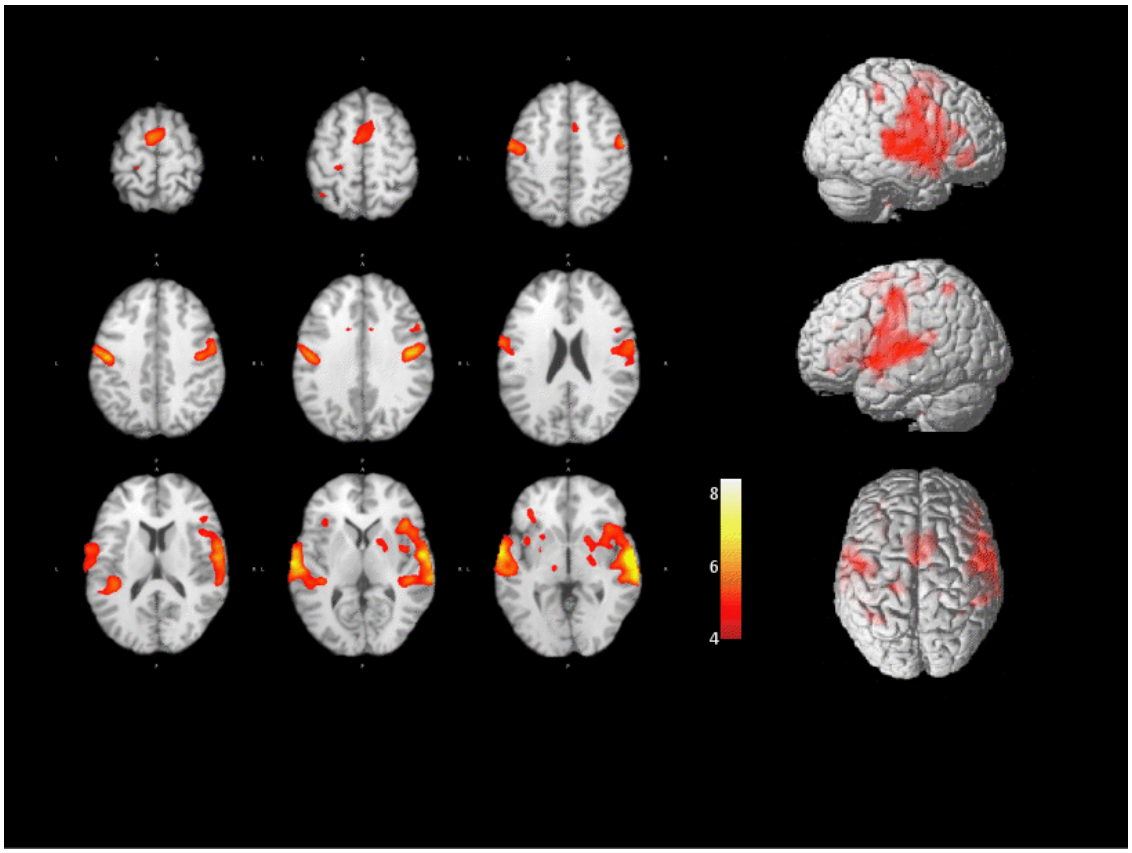


Figure 4.1 Contrast of BOLD activations in Vocalize vs. Rest conditions.

In a second analysis, we contrasted Shift vs. No Shift vocalization trials to target the regions involved specifically in the pitch-shift reflex. A fixed effects analysis contrasting Shift-No Shift conditions showed activations in SMA as well as bilateral STG, M1, PMC, and cerebellum. The same contrast was then performed using a random effects analysis, revealing bilateral activation in the superior temporal gyrus and a small cluster of activation within right IFG (Figure 4.2). Within STG a single cluster emerged with two areas of peak activation. The first was an anterior peak within the Sylvian fissure corresponding to the primary auditory cortex. A larger posterior peak was also found in the posterior STG near the Sylvian parietal temporal area (Spt), a region that recent studies have implicated in sensorimotor integration of the speech system (Hickok, Houde, & Rong, 2011).

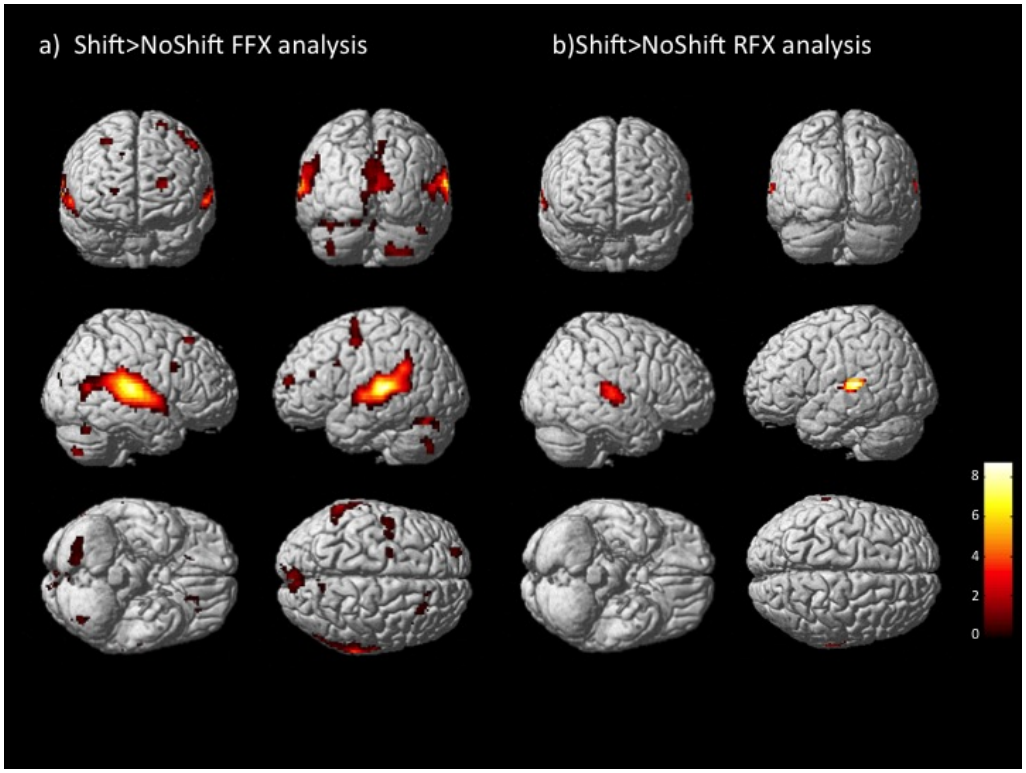


Figure 4.2 BOLD responses to the Shift – NoShift contrast. (a) Fixed effects analysis of the Shift- NoShift contrast ($p<0.05$, FWE). (b) Random effects analysis of the Shift – NoShift contrast ($p<0.001$).

4.2 Results of current fMRI study

While the previous study sought to define the nodes of activation specific to the pitch-shift reflex, the aim of the present study was to further specify the areas involved during a correct compensatory response as well as those involved in erroneous responses to pitch perturbations. As each fMRI trial of pitch-shift vocalization conditions consisted of 4-5 pitch shifts, the trials were divided up into 3 categories –compensatory (4/5 opposing responses), following (4/5 following responses), and mixed trials. While the resulting compensatory responses outnumbered the following responses as expected, approximately half of the fMRI trials were classified as mixed responses.

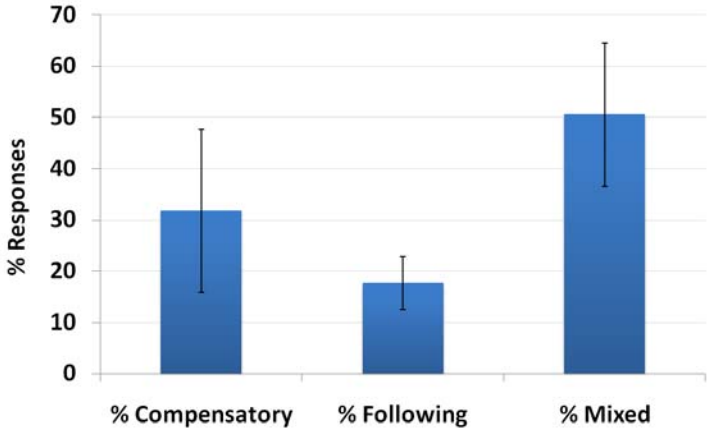


Figure 4.3 Percentages of fMRI trials classified compensatory, following, or mixed response trials.

A contrast analysis of Compensatory>Following responses revealed stronger activations in bilateral insula, left STG, right IFG and right cerebellum (Figure 4.4; p<.001 uncorrected). The increased BOLD activation in these areas during the production of a typical compensatory response suggests that these regions are of particular importance for error detection and correction. These results are discussed below in relation to models of speech production. In addition, we performed a separate contrast analysis comparing Following>Compensatory responses. Results from this contrast revealed activations in the right inferior parietal cortex, right superior occipital gyrus, right precuneus, posterior cingulate, middle cingulate, bilateral superior frontal gyrus, left middle frontal gyrus, and middle orbital gyrus (Figure 4.5; p<.001 uncorrected). These results are more difficult to interpret for two reasons. First, by narrowing our contrasts strictly to compensatory and following responses, we

excluded trials which demonstrated mixed responses, thereby further lowering the statistical power. Second, this is the first attempt to characterize the neural substrates involved in erroneous responses to the pitch-shift task. As a result, a precise explanation as to why following responses often occur in pitch-perturbation tasks has yet to be established. In consideration of these limitations, we provide a cautious interpretation of these results below and suggest a lack of attention to the task as a possible contribution.

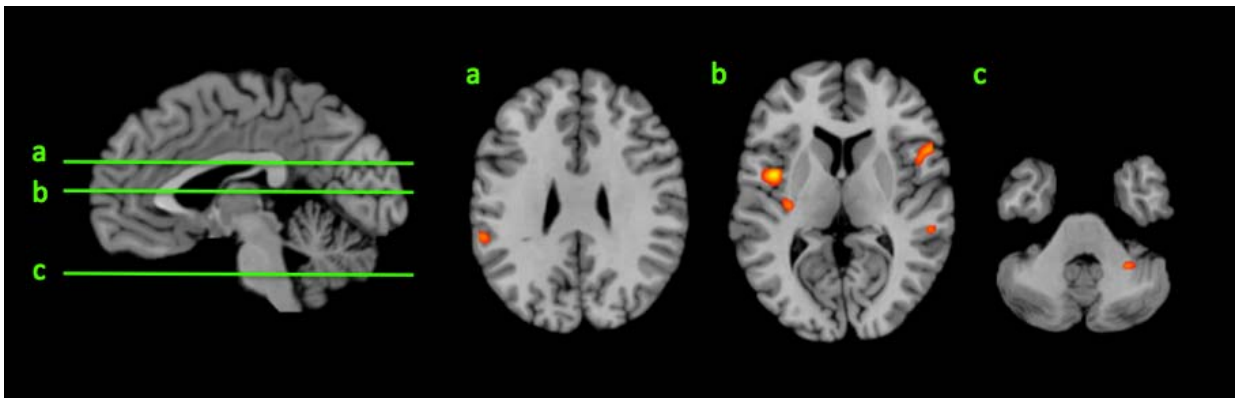


Figure 4.4 Contrast analyses of BOLD activations in Compensatory vs. Following responses to the pitch-shifted auditory feedback. (p<.001, uncorrected)

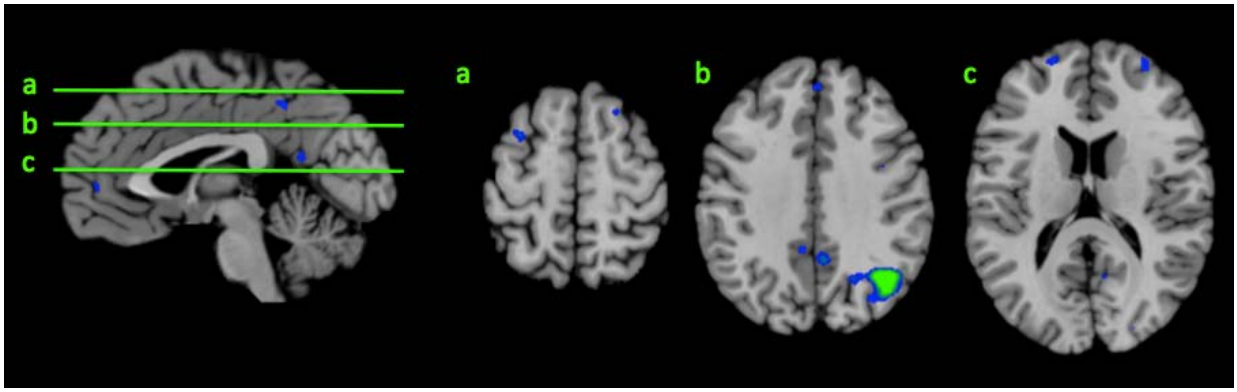


Figure 4.5 Contrast analyses of BOLD activations in Following vs. Compensatory responses to the pitch-shifted auditory feedback. (p<.001, uncorrected)

4.3. Previous TMS results

Our previous data also revealed an interaction between the presentation of a pitch shift stimulus and transcranial magnetic stimulation of the laryngeal motor cortex (LMC) during a pitch-shift reflex, using the methods described in the previous section. Subjects sustained an /a/ phonation while receiving randomized upward and downward pitch-shifted stimuli. During vocalization, single pulse TMS was delivered to the LMC at 0, 10, 20, 30, 40, and 50 ms following the onset of the pitch-shift stimulus. The magnitude of voice F0 responses to upward and downward shifted stimuli revealed interaction effects at a delay time of 30 ms when paired with irTMS. TMS coupled with an upward pitch-shift produced a decrease in response magnitude at 30-40ms after presentation of shifted feedback, while downward pitch-shift paired with TMS led to an increase in response magnitude (Figure 4.6).

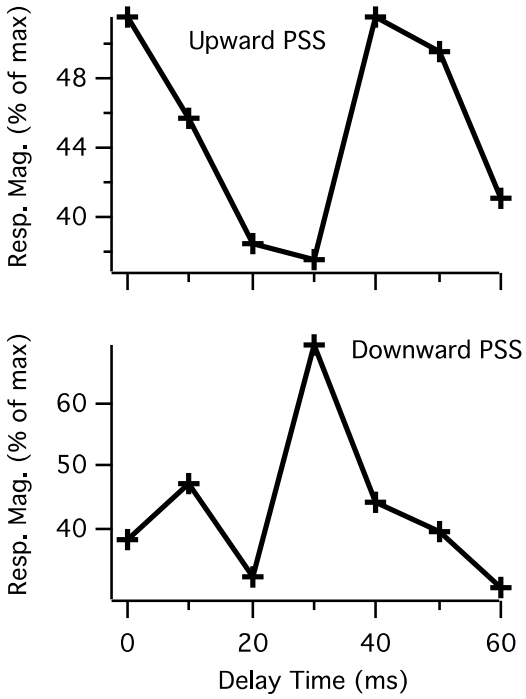


Figure 4.6 Changes in F0 response magnitude following upward (above) and downward (below) pitch-shifted feedback during TMS.

4.4 Results of Current TMS Study

The present study sought to control for generalized effects of TMS stimulation of the cortex. We targeted vertex as a non-speech cortical area to determine whether or not the effects observed in the previous study were specific to areas involved in speech production. The previous study demonstrated that TMS of the LMC during non-shifted vocalization led to an increase in voice F0 along the magnitude of 30-40 cents (Figure 4.7). However, stimulation of vertex during vocalization led to no apparent changes in voice F0 (Figure 4.8). Although further statistical analysis is needed, this suggests that the effects of TMS observed in the previous study were in fact specific to the speech system.

Figure 4.7 Effects of TMS stimulation of M1 during vocalization (Larson, et al., 2010)

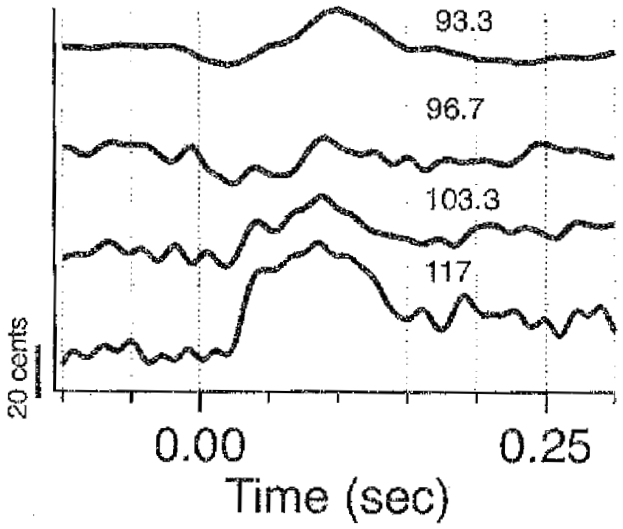
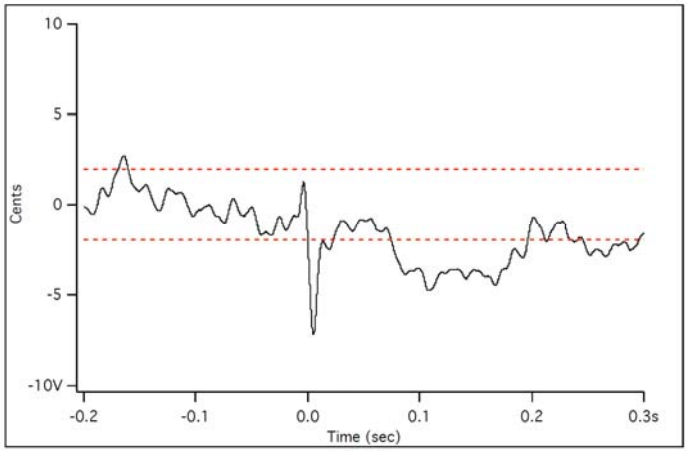


Figure 4.8 Effects of TMS stimulation of vertex during vocalization



4.5. Discussion

The discussion that follows addresses both the previous fMRI and TMS findings from our group and the current ones from this study. The purpose of this study was to (1) use BOLD fMRI to compare the cortical regions active during following versus compensatory responses in the pitch shift reflex and (2) to determine if interaction effects observed in the previous study reflect a generalized effect of TMS or an effect specific to speech areas involved in the pitch-shift reflex. The results of the paper allowed us to address both aims. In brief, fMRI data showed differing patterns of activation when comparing compensatory and following responses to pitch-

shifted auditory feedback. Secondly, while our previous work has shown that TMS of the LMC leads to changes in voice F0 both with and without perturbation, TMS of a non-speech control site (vertex) led to no systematic change in voice F0 during vocalization. Below these results are discussed in detail.

Part 1 – fMRI

The fMRI work leading up to the current study provided strong support for data from previous studies on the brain regions associated with vocalization (Brown, et al., 2009). As well, our work on the regional contributions to the pitch-shift reflex are generally concordant with other sensory perturbation changes during speech (Tourville, et al., 2008). In particular, an examination of fMRI data using a fixed effects analysis suggests the involvement of bilateral STG, PMC, and IFG in responding to auditory perturbations. The importance of the bilateral STG activations associated with the behavioral response to an unexpected pitch shift during vocalization is highlighted by the fact that this is the only region that survived the more conservative statistical approach that utilized a random effects model. To understand these brain effects it is important to provide a model based framework of speech/voice motor control.

Several models have been proposed to describe integration of the auditory and motor systems involved in speech. Early feedback models such as that of Fairbanks (1954) proposed a speech system that sought to minimize discrepancies between representations of predicted sensory outcomes and actual sensory feedback. Since then, numerous researchers have studied the SMI of the speech system in terms of both feedback and feedforward mechanisms of control. As described earlier, these feedforward mechanisms of error detection have been described in terms of efference copy– a process through which a copy of the motor output command is sent to the sensory systems to "cancel" out the sensory consequences of the given action. Thus, during

the execution of a motor action, there is a corresponding suppression of the relevant sensory areas in the brain – an effect referred to as motor induced suppression. For instance, research using evoked potential methods (which have low spatial resolution) has demonstrated that when sensory feedback of self-produced speech matches the predicted sensory outcome, there is suppression of activity in the auditory cortex which subsumes STG (Houde, Nagarajan, Sekihara, & Merzenich, 2002; Martikainen, Kaneko, & Hari, 2005). However, if a mismatch occurs between sensory feedback and the predicted sensory outcome of a motor action, an error signal is produced to correct the erroneous response. In this case, there is a release from suppression. As the difference between expected and actual production increases, the suppression systemically decreases. This continues until the feedback becomes an alien voice, at which point there is no longer any suppression of activity in auditory cortex (Liu, Behroozmand, & Larson, 2010).

The Directions Into Velocities of Articulators (DIVA) model provides both computational and empirical data to interpret findings from our work on perturbation. The DIVA model provides a detailed computational framework of the neural substrates underlying speech acquisition and production. According to this model, speech production begins with the activation of speech sound maps in the left posterior IFG and proceeds to activate both feedback and feedforward control systems. Both the speech sound maps in pIFG and initiation maps in SMA activate articulator velocity and position maps located in the primary motor cortex, leading to the initial production of speech. Within the feedback control system, auditory target, state, and error maps are located in both Heschl's gyrus and pSTG. Mismatches between predicted and actual sensory outcomes of self-produced speech are detected by auditory error cells, located in the posterior superior temporal gyrus (pSTG) and Heschl's gyrus (HG). Once an

error is detected, these error cells then activate feedback control maps in the right PMC which then relay commands to the motor cortex to correct the motor output.

As the pitch-shift paradigm produces a mismatch between predicted feedback of vocal pitch and the perceived auditory feedback, we expected that our results would map well onto the components of the DIVA model relevant to auditory feedback control. Results from our fixed effects analysis match closely with the activations reported by Tourville (2008) with activations in bilateral STG, PMC, and IFG. Based on the DIVA model, the observed activations in STG may be interpreted as corresponding to the aforementioned auditory error maps, while activations in the PMC and IFG likely have a role in sending a corrective command to the motor cortex. Contrary to our expectations and previous literature, M1 did not emerge as a significant area of activation when comparing Shift-No Shift. However, this result may have been due to the lack of power accompanied with sparse sampling. As noted above, the random effect analysis of our data revealed activation only in STG. However, within that cluster, there were two peaks in the BOLD signal – an anterior peak near Heschl's gyrus (primary auditory cortex) and well as a peak in pSTG. These two locations correspond to the areas described by the DIVA model as comprising auditory target, state, and error maps, which function to detect errors in sensory feedback of self-produced speech.

While these results appear to coincide with previous research on auditory perturbations and fit within the current models of speech production and error detection, these studies have failed to account for differing responses to perturbation paradigms. Two classes of responses to pitch perturbations have been previously documented by Burnett (1998). These consist of compensatory (or "opposing") responses, which indicate that a mismatch between predicted and expected auditory feedback has been detected and corrected for, and following responses, which

indicate that the online mechanism if error detection and correction has gone awry. While compensatory responses outnumbered following responses as expected, approximately half of the fMRI trials were mixed. The need to analyze these results separately is therefore crucial in ascertaining which activations are specific to neural mechanisms of error detection and correction.

One possible explanation for the differences in F0 response and BOLD activation patterns could be related to differences in attention and orientation to the task. Two separate pathways of the attention network have been proposed – a dorsal network and ventral network (M. D. Fox, Corbetta, Snyder, Vincent, & Raichle, 2006). The dorsal pathway is bilateral, encompassing the inferior parietal cortex (IPC) and frontal eye field (FEF) and has been linked to top-down orientation of attention. The ventral pathway, on the other hand, is a right lateralized network involved in reorienting attention in response unexpected changes sensory stimuli. The ventral pathway is characterized by nodes of activation in temporal-parietal junction (TPJ) and ventral frontal cortex (VFC). Increased BOLD response in the bilateral dorsal pathway has been associated with shifting attention to anticipated changes in sensory stimuli. Conversely, activation along the ventral pathway has been linked to re-orienting attention in response to unexpected shifts in sensory stimuli.

When examining the contrasts comparing following and compensatory responses, differences in activation are seen in several nodes involved in these attention networks. For instance we observed higher activation in right IFG in compensatory trials and higher activation in right IPC and superior occipital gyrus (SOG) in following trials. Although these data are far from conclusive, it is possible that these different response classes result from activation of different attention pathways whereby following responses reflect a top-down mechanism of the

dorsal pathway and compensatory responses reflect a ventral pathway response to unexpected change in sensory feedback. Another possibility is that for some trials subjects simply weren't attending or oriented to the task while in the scanner. These data suggest that following and compensatory responses to the pitch-shift paradigm reflect different underlying neural responses in the brain. However, another study is needed to fully address this issue. In particular, the lack of a fully event-related design limited the ability to analyze pure compensatory versus following responses and also resulted in a substantial drop in power.

Part 2 - TMS

Our second goal in this study was stimulate a non-speech control site (vertex) to examine whether the effects of TMS stimulation of M1 (LMC) were specific to locations encompassed by the speech system or whether they were due to a generalized effect of TMS. Using our innovative use of single pulse TMS to assess timing of the response between regions of the brain, we predicted in our previous study that stimulation of M1 during a pitch-shift task would produce and interaction effect, leading to a change in the magnitude of F0 response at the time of M1 activation. As we have reported, an interaction was found for both upward and downward pitch-shift conditions at 30ms post stimulus onset, with upward pitch-shift trials displaying a decrease in response magnitude and downward pitch-shift trials displaying an increase in response magnitude at that time.

We have interpreted these previous findings in the following way. Vocal pitch is regulated in part by the tensing and relaxing of the vocal folds. Tensing of the vocal folds produces a higher pitch (or increase in F0), while relaxation of the vocal folds produces a lower pitch (or decrease in F0). For upward pitch-shift trials voice F0 is reflexively lowered, and the

vocal folds relaxed, to compensate for the change in feedback. However, as the subject undergoes TMS of the LMC, the induced action potentials in M1-larynx result in a subsequent constriction of the laryngeal muscles. This prevents the vocal folds from reaching their initial degree of compensatory relaxation in response to upward pitch-shifts. Thus TMS during upward pitch-shift trials produces a decrease in the magnitude of response. In contrast, responses to downward shifted trials are characterized by tensing of the vocal folds and subsequent increase in voice F0. In this case, TMS of LMC produces an excitation of laryngeal muscles in addition to the excitation already produced in response to the downward pitch-shift. Thus, TMS during upward pitch-shift trials results in an increase in the magnitude of response.

The earliest latency EMG of a cricothyroid contraction in response to auditory stimulation was recorded at approximately 50ms (Sapir, McClean, & Larson, 1983). Meanwhile, the timing between activation of LMC (M1) and contractions of the chricothyroid muscle in the larynx has been estimated to be 15 ms (R. M. Rodel et al., 2004). We therefore hypothesized that the LMC was active approximately 35ms after stimulus onset. However, the interaction effect observed with TMS indicates that this time estimate may be closer to 30ms post-onset. However, the findings above are only valid if non-vocal regions of the brain show no interaction between voice production and TMS. The absence of any change in F0 during simultaneous vocalization and stimulation of vertex provides important validation of our method in that stimulation of areas unrelated to speech production do not alter vocal production. This allows for greater clarity in interpreting the results of TMS stimulation of M1 during vocalization with and without pitch-shifted feedback.

CHAPTER 5: CONCLUSIONS

The first aim of this study was to characterize the neural mechanisms underlying audio-vocal reflexes in healthy subjects, specifically in differentiating responses to pitch perturbations. Our previous results have implicated the roles of A1 and pSTG in detecting the mismatch between expected auditory feedback of vocal pitch and perceived auditory feedback. In addition, PMC and IFG are likely involved in the corrective mechanism responsible for producing compensatory changes in voice F0. In the present study, we were also able to demonstrate differences in BOLD activation patterns accompanied by following versus compensatory responses to the pitch-shift task, which we suggest might be attributed to differences in attention to the task.

In the previous study, we also sought to determine the timing of LMC activation in the pitch-shift reflex using a novel application of transcranial magnetic stimulation. An interaction was observed between TMS and presentation of the pitch shift stimulus 30ms after the onset of the pitch perturbation. The present study was aimed at using a non-speech control site to determine whether the effects observed in our previous data were site specific or due to a main effect of TMS stimulation. As stimulation of vertex led to no change in F0 during vocalization, it can be concluded that our results were in fact specific to the location. By incorporating these findings into previous research a temporal model of the pitch-shift reflex can be developed. As it stands, the response begins 12 ms after stimulus onset with activation of the auditory cortex. From this point, it proceeds to activation of the laryngeal motor cortex 30 ms after stimulus onset. At 50ms, there is then a contraction of the crichothyroid muscle in the larynx, after which a compensatory change in F0 is produced.

CHAPTER 6: IMPLICATIONS OF RESEARCH

The observed differences in behavioral and neural responses to the pitch-shift paradigm are important to consider for future research endeavors. Since they appear to be indicative of different neural processes, finding ways to control for or minimize erroneous response to pitch-perturbation will be important in future studies. An approach to this might be to use additional visual or auditory cues prior to vocalization trials to assure the subject's focus is on the task. Importantly though, it calls for a better understanding of why these two different responses emerge, for instance whether it is in fact due to differences in attention or some other breakdown in the SMI process. In future studies we plan to use an event related design in which each fMRI volume collected will reflect exactly one pitch perturbation and response. This will provide greater statistical power to fully examine these responses independently from another, should following responses once again be prevalent. Using this design, we also plan to conduct performance correlations of F0 responses with BOLD signal. This would allow us to again differentiate following and compensatory responses while establishing more concrete relationships between each response and the cortical areas involved in modulating them.

Defining the cortical areas involved in the audio-vocal reflex and establishing a temporal model of activity in those regions allows for a better understanding of the basic mechanisms of sensory voice control. Characterizing the roles of M1, PMC, IFG, and STG in voice production can, in the future, direct research on specific voice disorders in which the motor control of speech production is impaired. For instance, data from Dr. Larson's group have suggest that individuals with Parkinson's disease respond with greater compensation to pitch-shifted feedback than healthy controls, possibly indicating deficits in SMI and voice control. As such, we might expect to see a greater BOLD response to pitch-perturbations in STG, PMC, or IFG.

From there it may be possible to determine whether such differences might be localized to a specific region or attributed to a more general hyperactivity or the error correction system.

An in-depth investigation of the neural mechanisms involved in SMI of voice production would foster the development of more precise treatments for such disorders that focus on the sensorimotor integration of voice control rather than simply the motor control of speech. Ideally, the results of this study could also be used to further the investigation of general sensorimotor integration and contribute to our knowledge of other anomalies in motor control.

REFERENCES

Banziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*(3-4), 252-267.

Bauer, J. J., & Larson, C. R. (2003). Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: improvements in methodology for the pitch-shifting technique. *J Acoust Soc Am, 114*(2), 1048-1054.

Bauer, J. J., Mittal, J., Larson, C. R., & Hain, T. C. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: an automatic mechanism for stabilizing voice amplitude. *J Acoust Soc Am, 119*(4), 2363-2371.

Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2009). The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain Cogn, 70*(1), 31-41.

Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cereb Cortex, 18*(4), 837-845.

Bullmore, E., Brammer, M., Williams, S. C., Rabe-Hesketh, S., Janot, N., David, A., et al. (1996). Statistical methods of estimation and inference for functional MR image analysis. *Magn Reson Med, 35*(2), 261-277.

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *J Acoust Soc Am, 103*(6), 3153-3161.

Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice F0 responses to pitch-shifted voice feedback during English speech. *J Acoust Soc Am, 121*(2), 1157-1163.

Cooper, A. C., Humphreys, G. W., Hulleman, J., Praamstra, P., & Georgeson, M. (2004). Transcranial magnetic stimulation to right parietal cortex modifies the attentional blink. *Exp Brain Res, 155*(1), 24-29.

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *J Acoust Soc Am, 77*(6), 2142-2156.

Donath, T. M., Natke, U., & Kalveram, K. T. (2002). Effects of frequency-shifted auditory feedback on voice F0 contours in syllables. *J Acoust Soc Am, 111*(1 Pt 1), 357-366.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis and management.* St. Louis: Mosby.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *J Acoust Soc Am, 80*(2), 402-415.

Epstein, C. M., Fernandez-Beer, E., Weissman, J. D., & Matsuura, S. (1991). Cervical magnetic stimulation: the role of the neural foramen. *Neurology, 41*(5), 677-680.

Epstein, C. M., Lah, J. J., Meador, K., Weissman, J. D., Gaitan, L. E., & Dihenia, B. (1996). Optimum stimulus parameters for lateralized suppression of speech with magnetic brain stimulation. *Neurology, 47*(6), 1590-1593.

Epstein, C. M., Meador, K. J., Loring, D. W., Wright, R. J., Weissman, J. D., Sheppard, S., et al. (1999). Localization and characterization of speech arrest during transcranial magnetic stimulation. *Clinical Neurophysiology, 110*(6), 1073-1079.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., Peters, T. M. (1993). ***3D statistical neuroanatomical models from 305 MRI volumes.*** **Paper presented at the IEEE-Nuclear Science Symposium and Medical Imaging Conference.**

Fairbanks, G. (1954). Systematic research in experimental phonetics. I. A theory of the speech mechanism as a servosystem. ***J Speech Hear Disord, 19*****(2), 133-139.**

Fisher, R. J., Sharott, A., Kuhn, A. A., & Brown, P. (2004). Effects of combined cortical and acoustic stimuli on muscle activity. ***Exp Brain Res, 157*****(1), 1-9.**

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. ***Magn Reson Med, 33*****(5), 636-647.**

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. ***Proc Natl Acad Sci U S A, 103*****(26), 10046-10051.**

Fox, P. T., Narayana, S., Tandon, N., Sandoval, H., Fox, S. P., Kochunov, P., et al. (2004). Column-based model of electric field excitation of cerebral cortex. ***Hum Brain Mapp, 22*****(1), 1-14.**

Friston, K. (2002). Functional integration and inference in the brain. ***Prog Neurobiol, 68*****(2), 113-143.**

Friston, K. J., & Price, C. J. (2001). Generative models, brain function and neuroimaging. ***Scand J Psychol, 42*****(3), 167-177.**

Gosselin-Kessiby, N., Kalaska, J. F., & Messier, J. (2009). Evidence for a proprioception-based rapid on-line error correction mechanism for hand orientation during reaching movements in blind subjects. ***J Neurosci, 29*****(11), 3485-3496.**

Gracco, V. (1991). Sensorimotor Mechanisms in Speech Motor Control. In H. Peter, and Starkweather (Ed.), ***Speech Motor Control and Stuttering*** **(pp. 53-76). New York, New York: Excerpta Medica.**

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. ***Brain Lang, 96*****(3), 280-301.**

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., & Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. ***Exp Brain Res, 130*****(2), 133-141.**

Hammer, M. J., & Barlow, S. M. (2010). Laryngeal somatosensory deficits in Parkinson's disease: implications for speech respiratory and phonatory control. ***Exp Brain Res, 201*****(3), 401-409.**

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. ***Neuron, 69*****(3), 407-422.**

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. ***J Cogn Neurosci, 14*****(8), 1125-1138.**

Ito, T., Kimura, T., & Gomi, H. (2005). The motor cortex is involved in reflexive compensatory adjustment of speech articulation. ***Neuroreport, 16*****(16), 1791-1794.**

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. ***Neuroimage, 17*****(2), 825-841.**

Jennum, P., Friberg, L., Fuglsangfrederiksen, A., & Dam, M. (1994). Speech Localization Using Repetitive Transcranial Magnetic Stimulation. ***Neurology, 44*****(2), 269-273.**

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). The mirror-neuron system: a Bayesian perspective. *Neuroreport, 18*(6), 619-623.

Kochunov, P., Lancaster, J., Hardies, J., Thompson, P. M., Woods, R. P., Cody, J. D., et al. (2005). Mapping structural differences of the corpus callosum in individuals with 18q deletions using targetless regional spatial normalization. *Human Brain Mapping, 24*(4), 325-331.

Lancaster, J. L., Narayana, S., Wenzel, D., Luckemeyer, J., Roby, J., & Fox, P. (2004). Evaluation of an image-guided, robotically positioned transcranial magnetic stimulation system. *Hum Brain Mapp, 22*(4), 329-340.

Larson, C. L., Flagmeier, S., Fox, P., & Robin, D. A. (2010). Neural Mechanisms underlying vocal responses to perturbations in pitch auditory feedback. *Auditory Research Meeting*. The Acoustical Society of Japan, 40(4), 281-286.

Liu, H., Behroozmand, R., & Larson, C. R. (2010). Enhanced neural responses to self-triggered voice pitch feedback perturbations. *Neuroreport, 21*(7), 527-531.

Lo, Y. L., & Fook-Chong, S. (2004). Ipsilateral and contralateral motor inhibitory control in musical and vocalization tasks. *Exp Brain Res, 159*(2), 258-262.

Loucks, T. M., Poletto, C. J., Simonyan, K., Reynolds, C. L., & Ludlow, C. L. (2007). Human brain activation during phonation and exhalation: common volitional control for two upper airway functions. *Neuroimage, 36*(1), 131-143.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cereb Cortex, 15*(3), 299-302.

McRobbie, D. W., Moore, E.A., Graves, M.J., Prince, M.R. (2003). *MRI - From Picture to Proton*. Cambridge: Cambridge University Press.

Michelucci, R., Valzania, F., Passarelli, D., Santangelo, M., Rizzi, R., Buzzi, A. M., et al. (1994). Rapid-Rate Transcranial Magnetic Stimulation and Hemispheric Language Dominance - Usefulness and Safety in Epilepsy. *Neurology, 44*(9), 1697-1700.

Natke, U., Donath, T. M., & Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *J Acoust Soc Am, 113*(3), 1587-1593.

Natke, U., & Kalveram, K. T. (2001). Effects of frequency-shifted auditory feedback on fundamental frequency of long stressed and unstressed syllables. *J Speech Lang Hear Res, 44*(3), 577-584.

Pascualleone, A., Gates, J. R., & Dhuna, A. (1991). Induction of Speech Arrest and Counting Errors with Rapid-Rate Transcranial Magnetic Stimulation. *Neurology, 41*(5), 697-702.

Pelizzari, C. A., Chen, G. T., Spelbring, D. R., Weichselbaum, R. R., & Chen, C. T. (1989). Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *J Comput Assist Tomogr, 13*(1), 20-26.

Petersen, N. T., Pyndt, H. S., & Nielsen, J. B. (2003). Investigating human motor control by transcranial magnetic stimulation. *Exp Brain Res, 152*(1), 1-16.

Poizner, H., Feldman, A. G., Levin, M. F., Berkinblit, M. B., Hening, W. A., Patel, A., et al. (2000). The timing of arm-trunk coordination is deficient and vision-dependent in Parkinson's patients during reaching movements. *Exp Brain Res, 133*(3), 279-292.

Poline, J. B., Worsley, K. J., Evans, A. C., & Friston, K. J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage, 5*(2), 83-96.

Rodel, R. M., Olthoff, A., Tergau, F., Simonyan, K., Kraemer, D., Markus, H., et al. (2004). Human cortical motor representation of the larynx as assessed by transcranial magnetic stimulation (TMS). ***Laryngoscope, 114*****(5), 918-922.**

Rodel, R. M. W., Olthoff, A., Tergau, F., Simonyan, K., Kraemer, D., Markus, H., et al. (2004). Human cortical motor representation of the larynx as assessed by transcranial magnetic stimulation (TMS). ***Laryngoscope, 114*****(5), 918-922.**

Sapir, S., McClean, M. D., & Larson, C. R. (1983). Human laryngeal responses to auditory stimulation. ***J Acoust Soc Am, 73*****(1), 315-321.**

Sivasankar, M., Bauer, J. J., Babu, T., & Larson, C. R. (2005). Voice responses to changes in pitch of voice or tone auditory feedback. ***J Acoust Soc Am, 117*****(2), 850-857.**

Sommer, M. A., & Wurtz, R. H. (2002). A pathway in primate brain for internal monitoring of movements. ***Science, 296*****(5572), 1480-1482.**

Sommer, M. A., & Wurtz, R. H. (2008). Visual perception and corollary discharge. ***Perception, 37*****(3), 408-418.**

Stewart, L., Walsh, V., Frith, U., & Rothwell, J. (2001). TMS produces two dissociable types of speech disruption. ***Neuroimage, 13*****(6), S45-S45.**

Terao, Y., Ugawa, Y., Enomoto, H., Furubayashi, T., Shiio, Y., Machii, K., et al. (2001). Hemispheric lateralization in the cortical motor preparation for human vocalization. ***J Neurosci, 21*****(5), 1600-1609.**

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. ***Neuroimage, 39*****(3), 1429-1443.**

Tunik, E., Poizner, H., Adamovich, S. V., Levin, M. F., & Feldman, A. G. (2004). Deficits in adaptive upper limb control in response to trunk perturbations in Parkinson's disease. ***Exp Brain Res, 159*****(1), 23-32.**

Tzelepi, A., Yang, Q., & Kapoula, Z. (2005). The effect of transcranial magnetic stimulation on the latencies of vertical saccades. ***Exp Brain Res, 164*****(1), 67-77.**

Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. ***Journal of Cognitive Neuroscience, 16*****(6), 978-987.**

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. ***Science, 269*****(5232), 1880-1882.**

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. ***Neuroimage, 14*****(6), 1370-1386.**

Xu, Y., Larson, C. R., Bauer, J. J., & Hain, T. C. (2004). Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences. ***J Acoust Soc Am, 116*****(2), 1168-1178.**

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. ***Journal of Phonetics, 33*****(2), 159-197.**