UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Beyond Lateralization: Sensory-Motor Control of Speech Production.

Permalink
https://escholarship.org/uc/item/3kf6r01x

Author
Kort, Naomi Sophia

Publication Date
2014

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Beyond Lateralization: Sensory-Motor Control of Speech Production.

by

Naomi Sophia Kort

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

AND

UNIVERSITY OF CALIFORNIA, BERKELEY

[redacted]

Copyright 2014
by
Naomi Sophia Kort

Acknowledgements.

I am grateful to many people for their contributions to this work and support throughout my graduate studies.

I would like to express my deepest appreciation to my mentors, Srikantan Nagarajan and John Houde. Sri has been a great thesis advisor and John has been a fantastic mentor. Sri introduced me to neuroimaging and has encouraged my independent exploration of scientific ideas. John has guided and shaped my understanding of speech neuroscience. Together they have been an amazing mentoring team, helping me hone my critical thinking skills and grow as a scientist.

Many faculty members at UCSF have played an important role in my graduate career. My thesis committee, Robert Knight, Heidi Kirsch, and Edward Chang, provided valuable scientific insight. Heidi Kirsch, additionally, provided great mentorship on a clinical neuroimaging project not included in this thesis. Steve Lehman was a fantastic teaching mentor who showed me how to develop innovative and compelling curriculum.

My fellow graduate students, Julia Owen and Alexander Herman, have helped shape my thinking with our conversations.

Pablo Cuesta developed the functional connectivity analysis used in this thesis.

Several members of the Biomagnetic Imaging Lab have helped with my data collection and analysis. Susanne Honma, Danielle Mizuiri, Anne Findlay, and Maria Ventura all provided invaluable support for this work.

Leighton Hinkley, David Wipf, Carrie Niziolek, Zarinah Agnew, and Karuna Subramaniam, the post-docs in the lab, have helped to guide my scientific thinking through scientific discussions and technical guidance.

My parents, James Kort and Wilhelmina Smeets, my sister Rachel, and my brother Eric, gave me their complete confidence while they challenged me to improve.

My husband, Zachary Powers, has supported me with his love, patience, and encouragement.

Beyond Lateralization: Sensory-Motor Control of Speech Production.

Naomi Sophia Kort

The task of speaking is the translation of thoughts into an acoustic wave through the coordination of hundreds of muscles. Sensory-motor transformations are continuously occurring during ongoing speech as feedback is monitored and errors in production are rapidly recognized and corrected. While it is known that auditory feedback is used to monitor and maintain proper pitch production, the neural control of this important aspect of speech production remains poorly understood. Modulation of the timing and frequency of vocal pitch conveys important semantic and affective information during speaking. The goal of this thesis was to study sensory-motor control of speech production by examining the neural circuits involved in the auditory feedback control of pitch. In order to achieve this aim we used a combination of magnetoencephalography (MEG) and real-time pitch altered auditory feedback. To this end, we conducted a series of experiments to understand the real-time cortical monitoring of one's own speech production, the recognition of errors in auditory feedback, the cortical processing of these errors, and the motor (and subsequently acoustic) compensatory change. In this series of studies we have described the cortical networks and their mechanisms in monitoring auditory feedback during speech production. We have demonstrated that the cortical monitoring of the onset of a vocalization occurs with the suppression of auditory, cerebellar and frontal regions in both the left and right hemisphere. We showed this suppression is released when the perceived auditory feedback does not match the expected auditory feedback. We investigated the cortical networks involved in

monitoring ongoing productions for errors in feedback, and the behavioral and cortical response when an error is perceived. We have shown that while the cortical networks monitoring the onset of speech and mid-utterance are overlapping, that these networks have distinct timing, patterns of response, and anatomical locations. Importantly, both feedback monitoring and motor control of speech at onset and mid-utterance show changes in cortical dynamics in both hemispheres. Feedback control of pitch during a mid-utterance error has corresponding increased inter-hemispheric communication. The work in this thesis provides evidence that inter-hemispheric communication is important for the feedback control of vocal pitch production.

Table of Contents

List of Tables.

List of Figures.

Chapter 1.
Introduction.

The task of speaking is the translation of thoughts into an acoustic wave. The successful execution of this task requires the coordination of hundreds of muscles to rapidly construct strings of sounds within a constrained acoustic space. Despite the complexity of speaking, people are highly skilled speakers. Yet, due in part to the complexity of the motor task of speaking, the neuroscience of how speech is produced is largely unknown. Traditional theories of speech production posit a left hemisphere lateralization in speech production (Broca, 1861; Dronkers, 1996; Hickok et al., 2011). This emphasis on functional lateralization neglects the role of inter-hemispheric communication and the right hemisphere in speech production, despite neuroimaging studies consistently showing bilateral neural activity to speech perception and during speech production (Price, 2010). Recent work has disrupted this conventional thinking by showing sensory-motor transformations occur bilaterally for word repetition (Cogan et al., 2014). Yet sensory-motor transformations do not only occur in speech repetition tasks, but are continuously occurring during ongoing speech as feedback is monitored and errors in production are rapidly recognized and corrected.

The act of vocalization is necessarily accompanied by concurrent sensory consequences-somatosensory feedback associated with the movement of articulators, and auditory feedback resulting from the movement. It is not surprising, therefore, that speakers monitor their sound output, and that this auditory feedback exerts a powerful influence on their speech. Yet, the role of auditory feedback during speech production is complex since auditory feedback is noisy and delayed. Indeed, the motor skill of speaking is very difficult to acquire without auditory feedback, and, once acquired, the skill is gradually lost in the absence of auditory feedback (Cowie et al.,

1982). While speech can continue in the absence of auditory feedback such as in the presence of very loud masking noise (Lane and Tranel, 1971), the control of the fundamental frequency (f0), or pitch, of speech is rapidly lost in the absence of auditory feedback (Lane and Webster, 1991). These findings demonstrate that pitch, along with other suprasegmental features, requires aural monitoring.

When auditory feedback is present, its alteration can have immediate effects on ongoing production. It has long been known, for example, that delaying auditory feedback can immediately render a speaker disfluent (Lee, 1950; Yates, 1963). More recently, experiments have altered specific features of auditory feedback, and the responses of speakers have been particularly revealing. In response to brief perturbations of the pitch, loudness, and formant frequencies of their auditory feedback, speakers will make quick adjustments to their speech that reduce the perceived effect of the perturbations on their auditory feedback (Chang-Yit et al., 1975; Houde and Jordan, 2002; 1998; Lane and Tranel, 1971; Lombard, 1911). These experiments, in which a feedback perturbation elicits a quick compensatory response, demonstrate the existence of speech sensorimotor pathways in the CNS that convey corrective information from auditory areas to speech motor areas during ongoing speaking. Behavioral experiments have further shown that auditory feedback is important for online control of pitch both in words and sentences. Altered pitch feedback on the first syllable of a nonsense word impacts the pitch of the second syllable, even when the first syllable is short and unstressed (Donath et al., 2002; Natke and Kalveram, 2001). Compensation to pitch-altered feedback influencing either the stress in a sentence (Patel et al., 2011) or the form of the sentence (Chen et al., 2007) has been observed. These studies have shown the importance of auditory feedback in controlling pitch both within a syllable, and on a suprasegmental level. This behavior may be used to compensate for disturbances in output pitch known to arise from a number of natural

sources, including an error in the complex coordination of vocal fold tension (Lane and Webster, 1991), aerodynamic instability, and even heartbeat (Orlikoff and Baken, 1989). The rapid compensation to altered pitch feedback occurs both in continuous speech (sentence production) and during single vowel phonation. This is not surprising given that phonation is an important part of speech. While the pitch perturbation response has been well characterized in behavioral studies, very little is known about the neural substrate of these sensorimotor pathways.

Until recently, the study of the neural circuitry monitoring self-produced speech has primarily focused on auditory cortex during correct (unaltered) vocalization. Work in non-human primates found that the majority of call-responsive neurons were inhibited during phonation (Eliades and Wang, 2008; 2005; 2002; Muller-Preuss and Ploog, 1981). Extensive work has been done to study this suppression effect in humans (Chang et al., 2013; Curio et al., 2000; Flinker et al., 2010; Greenlee et al., 2011; Houde et al., 2002; Ventura et al., 2009). Studies using magnetoencephalography (MEG) in humans similarly found suppressed neural activity in auditory areas during self-produced speech compared to the neural activity while listening to the playback of recorded speech (Curio et al., 2000; Houde et al., 2002). This effect has been termed speaking-induced suppression (SIS). SIS is a specific example of the broader phenomenon of motor-induced suppression (MIS) (Aliu et al., 2009), where sensory responses to stimuli triggered by self-initiated motor act are suppressed. However, in these studies, it was difficult to localize the SIS effect to specific areas of auditory cortex. Better localization of SIS has been seen in studies based on intracranial recording (ECoG) in neurosurgery patients. One study found that the SIS response varied not only across adjacent electrodes separated only by millimeters, but also across frequency bands (Greenlee et al., 2011). Some electrodes that showed SIS with a low frequency, evoked response either did not show SIS or showed anti-SIS with a high gamma band induced analysis. Another ECoG study focusing on responses in the left hemisphere found SIS primarily

in electrodes clustered in posterior superior temporal cortex (Chang et al., 2013). However, the spatial coverage of ECoG is limited, and, as yet, no studies to date have examined more completely the spatial distribution of SIS along the speech sensorimotor pathways or across frequency bands.

Monitoring feedback to confirm that speech motor acts give rise to the expected auditory outputs (resulting in SIS) is only one important role of the speech sensorimotor pathways. When feedback is altered and mismatches expectations, these pathways take on the additional role of conveying the mismatch to motor areas and generating a compensatory production change. What are the neural correlates of this process? Several SIS studies have showed that altering feedback at speech onset reduces SIS (Behroozmand and Larson, 2011; Heinks-Maldonado et al., 2006; Houde et al., 2002). These studies were conducted in sensor space and so included sources throughout and beyond the temporal lobe. Therefore, neither of these studies can address if the reduction of SIS is a true reduction of suppression by the increased activity of sources during speaking or is instead the superposition of suppression and enhancement. Agnew and colleagues, using fMRI, have show that aspects of SIS are independent of details of the acoustic aspects of speech, for they showed that SIS occurs in anterior superior temporal gyrus even when subjects are mouthing, but not overtly producing speech during sentence reading while hearing a different listener produce the sentence (Agnew et al., 2013). The Agnew and colleagues study used neural responses collected over the entire time of reading a sentence, so do not necessarily reflect the rapid (200ms window following vocalization onset) suppression associated with the onset of a vocalization. These questions can be answered by looking at the sources driving SIS and how the response during speaking changes with altered auditory feedback.

Feedback monitoring occurs not only at speech onset, but also during ongoing speech.

There is strong behavioral evidence for two types of monitoring systems- one at the onset of vocalizations and a distinct feedback monitoring system during ongoing vocalization. Trained singers are able to actively inhibit compensatory behavior when a sudden alteration in pitch occurs mid-utterance (Zarate and Zatorre, 2008), but are unable to inhibit compensation when the alteration occurs at speech onset (Keough et al., 2013). Comparing cortical responses to auditory feedback at speech onset and during mid-utterance can directly address the question of how these networks overlap. A few recent studies have looked at responses to feedback alterations during ongoing speech to pitch perturbations of auditory feedback using EEG (Behroozmand et al., 2009; Behroozmand and Larson, 2011; Behroozmand et al., 2011b). These studies found that perturbations of ongoing vocal feedback evoked larger responses than did perturbations passively heard during the subsequent playback of feedback. We term this effect speech perturbation response enhancement (SPRE). Although EEG studies to date have not been able to localize SPRE to particular brain areas, in two recent ECoG studies the spatial distribution of SPRE was mapped (Chang et al., 2013; Greenlee et al., 2013). One study from our group looked at high gamma responses to pitch-altered feedback in the left hemisphere, finding SPRE responses clustered in ventral premotor cortex and posterior superior temporal cortex including inferior parietal cortex (Chang et al., 2013). A second study, using ECoG, found enhanced evoked and high gamma responses in both left and right mid-to-anterior superior temporal gyrus (Greenlee et al., 2013). Coverage limitations of ECoG restrict the analysis to the individual subject's placement of the grid electrodes. Furthermore, since each patient's grid is uniquely placed ECoG studies cannot easily compare results across subjects, or across hemispheres within a subject.

The relationship between feedback monitoring at speech onset and mid-utterance- SIS and SPRE- is largely unexplored. A recent study in marmosets found auditory neurons that show

suppression at the onset of vocalization are more likely to have an enhanced response to altered auditory feedback at vocalization onset, suggesting a direct link between the mechanisms suppressing self-produced speech to those recognizing errors in self-produced speech (Eliades and Wang, 2008). Comparatively with these findings, an ECoG study of SPRE found only a small number of electrodes displaying both SIS and SPRE, while the majority of electrodes which preferentially display one or the other (Chang et al., 2013).

The goal of this thesis was to study sensory-motor control of speech production by examining the neural circuits involved in the auditory feedback control of pitch. In order to achieve this aim we used a combination of magnetoencephalographic imaging (MEG) and real-time pitch altered auditory feedback. In the first study we examined the evoked cortical responses at both speech onset (to examine SIS) and during brief, unexpected shifts in the pitch of subjects' audio feedback (to examine SPRE) during the phonation of a single vowel and contrasted the networks between these two events. In the second study, we investigated the spatial and frequency distribution of SIS on the entire cortex. We further tested the effect on SIS when the auditory feedback deviates both slightly and considerably from the expected acoustic consequence of speaking. In the third study, we sought to investigate which cortical regions are involved in voice pitch control during ongoing phonation, how these cortical responses to an error in pitch production evolve over time, how these cortical responses relate to behavioral responses, and the neural connectivity between nodes in the pitch production network.

Chapter 2.

A bilateral cortical network responds to pitch perturbations in speech feedback.

2.1 Introduction.

The goal of the present study was to examine the cortical evoked responses to both the onset of speech production and to an unexpected alteration in auditory pitch feedback mid-utterance. In particular, do areas in the speech sensorimotor pathways exhibit SPRE, and, if so, what is the time course of SPRE over these areas? Do the areas that exhibit SPRE also exhibit SIS? Do the areas that exhibit SPRE show correlations across subjects with behavior?

The present study used MEG to investigate the cortical neural responses at speech onset (to examine SIS) and during brief, unexpected shifts in the pitch of subjects' audio feedback (to examine SPRE) during the phonation of a single vowel. By using a single vowel utterance we were able to isolate phonation from additional (linguistic) aspects of speech to specifically study pitch production. In this study, we tested several hypotheses. First, given that SIS has been shown to be involved in auditory self-monitoring, we hypothesized that there would be spatial overlap between the monitoring role of SIS and the error recognition part of the SPRE network. Second, we hypothesized that SPRE would be seen to propagate through the speech sensorimotor network as the error is recognized and processed, and ultimately induce a compensatory response. Third, we hypothesized that cortical responses to the perturbation during speaking would be correlated with compensation across subjects.

2.2 Materials and Methods

2.2.1 Subjects

Eleven right-handed (4 female) English speaking volunteers with normal speech and hearing participated in this study. All participants gave their informed consent after procedures had been fully explained. The study was performed with the approval of the University of California, San Francisco Committee for Human Research.

2.2.2 MEG Recording

The task was completed during whole head MEG neural recording in awake subjects lying in the supine position. The MEG system (MISL, Coquitlam, British Columbia, Canada) consists of 275 axial gradiometers and was recorded with a sampling rate of 1200Hz. Three fiducial coils were placed on the nasion and left/right preauricular points to triangulate the position of the head relative to the MEG sensor array. In a separate session high resolution anatomical MRIs were obtained for each subject. The fiducial markers points were later co-registered with an anatomical MRI to generate head shape.

2.2.3 Experimental Design and Procedure

The subjects spoke into an MEG-compatible optical microphone and received auditory feedback through earplug earphones. They observed a projection monitor directly in their line of sight. The screen background was black. A trial began when three white dots appeared in the center of the screen. Each dot disappeared one by one to simulate a count-down (3-2-1). When all three dots disappeared and the screen was completely black, subjects were instructed to follow the instructions corresponding to the block- either speaking or listening. The trial was terminated

with the visual cue indicating the remaining trials before a break. A schematic of the experimental setup is shown in figure 2.1.

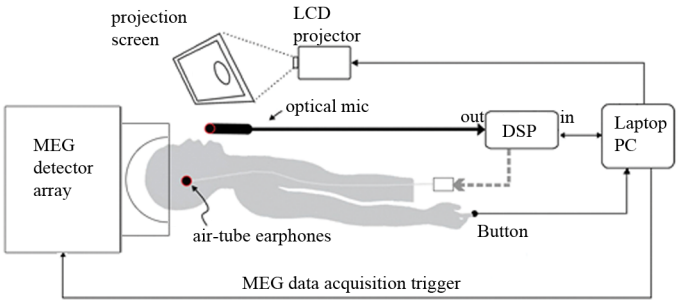

Figure 2.1. Schematic of experimental setup. Subjects are in the supine position in the MEG detector array. They speak into the optical microphone and receive auditory feedback through headphones. Their speech is passed through the Digital Signal Processor (DSP) both during altered and unaltered feedback. The laptop computer triggers the pitch-shifted stimulus at a jittered delay after speech onset.

The experiment consisted of 4 blocks of 74 trials each, with brief, self-paced breaks every 15 trials. The blocks were arranged in an alternation of conditions: blocks 1 and 3 were the Speaking Condition; blocks 2 and 4 were the Listening Condition. In the Speaking Condition, subjects were instructed to produce a sustained utterance of the vowel /a/ until the termination cue. During the phonation the subjects heard one 100-cent pitch perturbation lasting 400ms whose onset was jittered in time from speech onset. In each trial, the direction of the pitch shift was either positive (i.e., raising the perceived pitch) or negative (lowering the perceived pitch). Equal numbers of positive and negative perturbation trials were pseudorandomly distributed across the experiment. The jittered perturbation onset prevented the subject from anticipating the timing of the perturbation while the pseudorandom selection of a negative and positive pitch shift prevented the subject from anticipating the direction of the perturbation. In the Listening Condition, subjects saw the same visual prompts but only passively listened to the recording of their perturbed voice feedback obtained in the previous Speaking Condition block. The auditory input through earphones in both conditions was identical, providing a method of comparison to extract speaking-specific activity. The auditory input through the earphones was adjusted prior to

commencement of the experiment to a level that subjects reported their auditory feedback was the same as or slightly louder than expected.

2.2.4 Audio analysis: Pitch perturbation experiment

Speech recordings were analyzed for pitch throughout the utterance. Timing and magnitude of compensation was determined as follows. For each subject, the pitch contour of each perturbation type, +100 cent or -100 cent, were averaged together. Absolute frequency (hertz) was changed to cents peak response change from pre-perturbation baseline by: cents change = $100x[12x\log_2(\text{pitch response peak (Hz)}/\text{mean pitch frequency of pre-perturbation baseline (Hz)})]$. Mean percent compensation was calculated as -100*(cents change)/(cents of applied pitch shift). The negative sign makes the compensation a positive value. The pitch analysis was performed on each subject's single trial audio data. The audio data were sampled at 11025 Hz and both the microphone and feedback were recorded and analyzed. Each trial of the data was recorded in 32-sample frames and pitch was estimated for each of these frames using the standard autocorrelation method (Parsons, 1986). The resulting frame-by-frame pitch contour was then smoothed with a 20Hz, 5^{th} order, low pass Butterworth filter. Trials with erroneous pitch contours were removed. The largest possible pre-perturbation interval was used to establish a baseline, constrained by the minimum time between voice onset and perturbation onset, and the mean and standard deviation of this baseline interval was calculated. A subject's response onset was conservatively set to occur when the mean pitch time course deviates from the baseline by two standard deviations. The magnitude and onset of the compensation were determined for each subject and then averaged to create the grand-average compensation magnitude and onset. In this study, the data from ten subjects were included in this analysis. One subject was eliminated from the behavioral data analysis due to a corrupted file.

2.2.5 MEG data preprocessing

The MEG sensor data were marked at the speech onset and at the perturbation onset. Third gradient noise correction filters were applied to the data and the data were corrected for a DC offset based on the whole trial. Artifact rejection of abnormally large signals due to EMG, head movement, eye blinks or saccades was performed qualitatively through visual inspection and trials with artifacts were eliminated from the analysis.

2.2.6 Virtual sensor evoked analysis.

The time courses of source intensities at selected ROIs were computed using an adaptive spatial filtering technique using a signal bandwidth of 0-300 Hz (Bardouille et al., 2006; Oshino et al., 2007; Robinson and Vrba, 1999; Sekihara et al., 2004; Vrba and Robinson, 2001) with CTF software tools. A virtual sensor was created for each of the 4 defined locations for each task (speak or listen) and for each hemisphere (left or right) for every subject. The virtual channel data, i.e. source time course, were then filtered from 2 to 40 Hz. Each subject's data were normalized using the mean and standard deviation of the baseline activity of all sensors to compute a z-score. The z-scored results were used for averages and statistics. The latency and the magnitude of the peaks for each individual subject were calculated.

T-tests were computed over the eleven subjects comparing the speaking condition to the listening condition at each corresponding peak for both magnitude and latency differences. Multiple comparisons were corrected for using the 5% False Discovery Rate using the Benjamini and Hochberg method (Benjamini and Hochberg, 1995), which set significance at $p < 0.0096$. Since only a few planned comparisons were performed on each virtual sensor (only peak values were tested), all results $p<0.05$ are reported, and all uncorrected p-values given.

The ROI selection was conducted as follows. When the results from the altered auditory feedback fMRI studies are combined with other pertinent speech studies and theoretical models, four well-defined neural regions emerge as likely computational nodes in processing and responding to an error in auditory feedback. These four cortical regions became the regions of interest used for the virtual sensor analysis. These four ROI's are primary auditory cortex, the superior temporal gyrus/middle temporal gyrus (STG/MTG), ventral supramarginal gyrus/posterior superior temporal sulcus (vSMG/pSTS) and premotor cortex, and the supporting evidence for their role as computational nodes are summarized in Table 2.1 (Andersen et al., 1997; Buchsbaum et al., 2011; Fu et al., 2006; Gelfand and Bookheimer, 2003; Grefkes and Fink, 2005; Hickok et al., 2011; 2009; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009; Shadmehr and Krakauer, 2008; Tourville et al., 2008; Toyomura et al., 2007). The virtual sensor seeds were anatomically determined as the center of mass of the Brodmann areas (BA) corresponding with the aforementioned sensorimotor nodes: 41/42, 21/22, 40 and 6. The time course represents the activity in the region of interest approximately 1 centimeter from the center seed voxel due to spatial blur of the MEG source localization.

MNI Voxel Location	Talairach Voxel	Brodmann Area	Functional Area	References
[-54.3, -26.5, 11.6] [54.4, -26.7, 11.7]	[-51, -27, 13] [52, -27, 13]	41/42	Primary Auditory Cortex	pSTg: [-64,-30,14][11] PT: [-52,-34,16] [11] PT: [58,-28,12] [11] pSTG:[-62,-30,14][11] pdSTs: [58,-28,6] [11] pdSTS:[-60,-30,10][11] PT: [-62,-24,10] [11] STG/42: T: [53, -20, 9] [10] [57, −20, 15] [10] [−57, −20, 20] [10] Auditory Cortex[6,8] BA 41/42[6,12] STG/STS[8,12] Sensory System[7]
[-58, -19.8, -6] [58.9, -20.1, -6.1]	[-54, -21, -1] [56, -21, 0]	21/22	Superior Temporal Gyrus/ Middle Temporal Gyrus	STA: [52 -10 -2][2] STS: [-60 -30 3][1] STS: [63 -24 0][1] MTG/BA 21: T: [57, −10, −7][10] [57, -30, -2][10] [-53, -13, -7][10] [-53, -26, -7][10] STG/BA 22: T: [57, − 17, 9][10] [-57, -17, 4][10] adSTs: [56,-10,-4][11] Auditory Cortex[6,8] BA 21/22[6,12] STG/STS[8,12] Sensory System[7]
[-51.2, -43.3, 40.0] [51.8, -43.3, 40.4]	[-50, -41, 38] [52, -40, 38]	40	Supramarginal Gyrus/ posterior Superior Temporal Sulcus/ Inferior Parietal Lobule	IPS: [45, -36, 45] [1] IPS: [32, -42, 44][2] Spt: [-51, -42, 21][1] Supramarginal Gyrus: T: [-44, -40, 40][3] Spt: T: [-52, -43, 28][4] Spt: T: [-40, -32, 26][4] PO: [-44,-34,24][11] IPS[5] IPL, BA 39&40[6] Parietal Cortex[7,9] Spt[8]
[-46, 0, 35] [46.5, 0, 35]	[-45, 0, 33] [45, 0, 32]	6	Premotor Cortex	Premotor: [-52, 8, 38][2] Precentral Gyrus: T: [48 -4 36][3] [-34 2 38][3] Inferior Frontal Gyrus: T:[-50 10 36][3] Precentral Sulcus: [-51 -9 42][1] vPMC: [-48,0,30][11] Premotor Cortex[6,7,8,12] BA 6[6]

[1]Buchsbaum, 2011[2]Toyomura, 2007 [3]Gelfand, 2003 [4]Hickok, 2009 [5]Grefkes, 2005 [6]Rauschecker, 2009 [7]Shadmehr, 2008 [8]Hickok, 2011 [9]Andersen, 1997 [10]Fu, 2006 [11]Tourville, 2008 [12]Hickok, 2007

Table 2.1. The location of each virtual sensor is reported in both MNI and Talairach coordinates. The Brodmann area for which the voxel is the center of mass is listed and the corresponding functional area is given. Experimental and theoretical references are given that informed the choice of central voxel for the ROI analysis.

2.3 Results

2.3.1 Auditory pitch perturbations induce compensation.

The subjects responded to the pitch-shifted auditory feedback by adjusting the pitch of their voice to oppose the direction of the perturbation. The grand average across subjects is shown in figure 2.2. The average compensation to the perturbation across all subjects was 18.98% (range: 8.81% to 41.09%). The compensation for all subjects had a mean onset of 210.9ms (standard error = 19.87) after the perturbation onset and mean peak latency of 531.8ms (standard error = 14.07). The latency of the compensation onset was determined for each subject as the point in which the pitch contour crossed the two standard deviation line determined by the baseline period, and then averaged across subjects. Although the latency of the compensation onset in the averaged F0 contour seen in figure 2.2 appears to occur around 140ms, this is an inaccurate appearance of averaging across subjects. Compensation in response to the positive pitch shift and to the negative pitch shift were not statistically different either in magnitude ($p=0.744$) or in onset time ($p = 0.187$). Since the behavioral response to the positive and negative pitch shift trials was indistinguishable, all shifted trials were subsequently analyzed together.

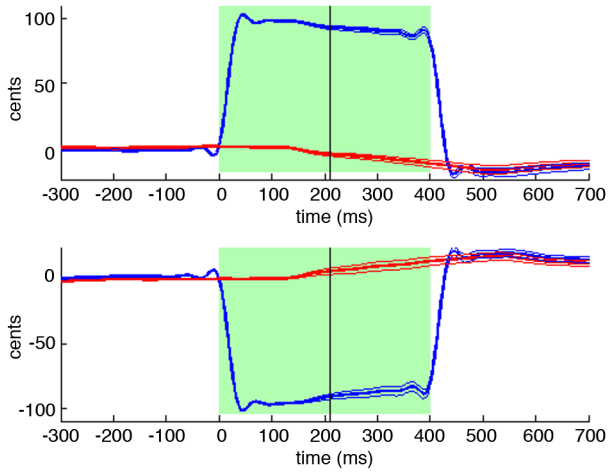

Figure 2.2. Vocal responses to the pitch shifts of audio feedback. Upper and lower plots show time courses of the average subject response to the +100 cent and -100 cent pitch perturbations, respectively. In each plot, time is relative to perturbation onset at 0ms, and pitch is expressed in cents relative to the mean pitch of the pre-perturbation period. Green regions indicate the duration of the pitch perturbation. Blue traces show the mean time course of feedback heard by subjects, with thick blue line is the grand-average over subjects and flanking thin blue lines are +/- standard errors. In a similar fashion, red traces show the mean time course of the pitch produced by subjects. The black vertical line denotes the mean time when the subjects' mean pitch time course deviates from the baseline by two standard deviations.

2.3.2 Cortical neural activity in response to speech onset.

Figure 2.3 shows the results for the four bilateral source time course reconstructions centered with zero at the time of speech onset in both the speaking and the listening conditions. In primary auditory cortex, we saw variations of the standard three-peak response to auditory input (M50, M100, M200). In the left primary auditory cortex this three peak pattern was clear in both the listening condition, (peaks at 30.9ms, 91.7ms, 188.6ms after speech onset), and in the speaking condition (peaks at 30.5ms, 92.9ms, 192.58ms after speech onset). In the right hemisphere, it was less clear that the pattern of response peaks fitted the standard three-peak auditory response pattern. In the listen condition, there were response peaks at 57.9ms, 84.8ms, 133.3ms, but there were also peaks at 31.7ms and 237.5ms. In the speaking condition, there were three clear peaks, but their correspondence to the standard auditory response pattern is

complicated by the pattern of latencies (35.6ms, 118.0ms, 235.3ms) where the M50 peak is early, and the M200 peak is late, and by the unusually large amplitude of the M50 peak.

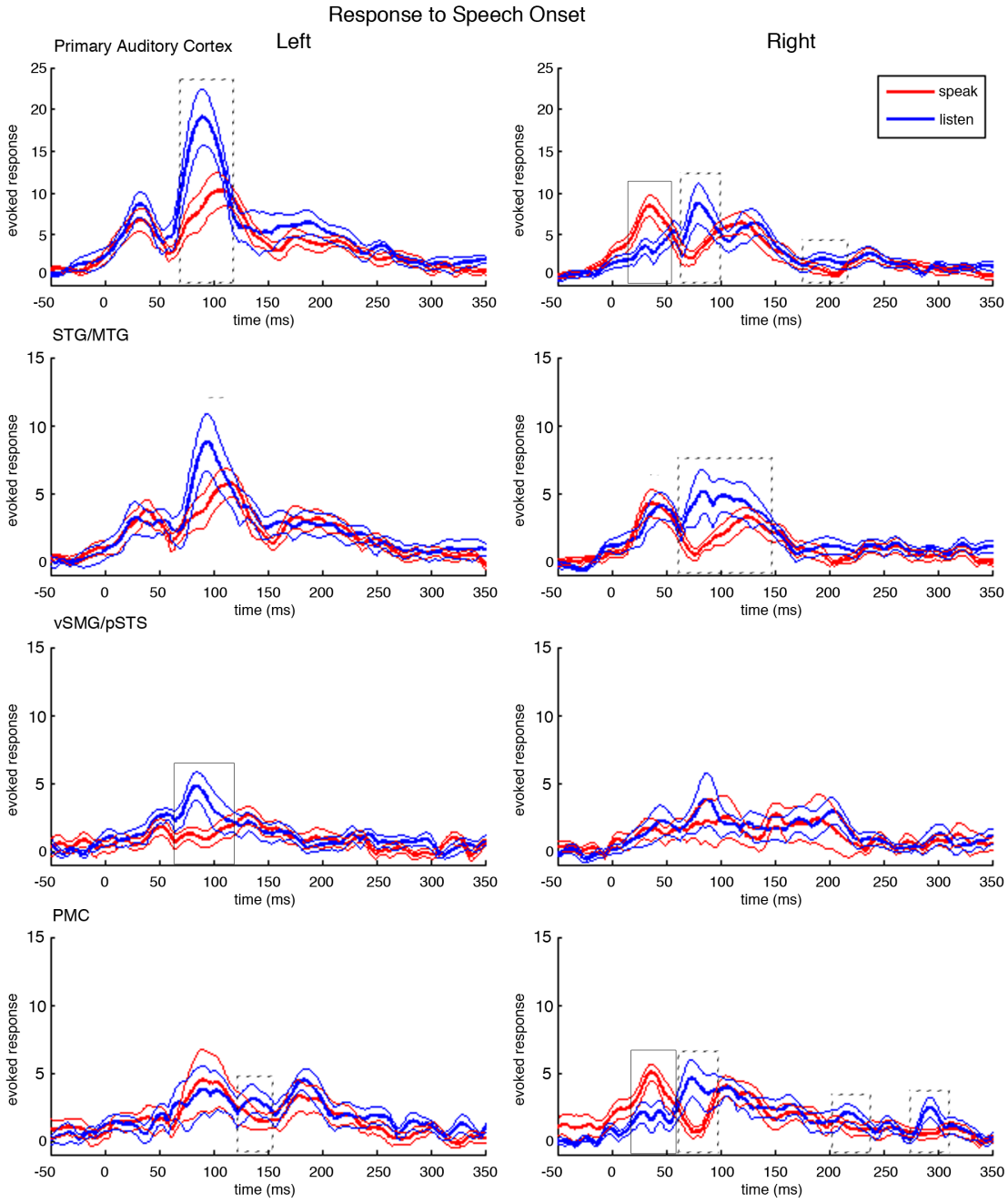

Figure 2.3. Neural responses to the onset of speech in both the speaking and the listening condition. In each plot, time is relative to the onset of speech at 0ms. Red traces show the mean cortical responses at the corresponding anatomical location across all subjects in the speaking condition. The thick red line is the grand-average over subjects and the flanking thin red lines are the standard error bars. Blue traces show the mean cortical responses in the listening condition. The thick blue line is the grand-average over subjects and the thin blue lines are the standard error bars. Solid boxes surround peaks with a significant magnitude difference between speaking and listening, $p < 0.0096$, dotted boxes surround peaks with $p<0.05$. Similarly, solid lines mark significantly different latencies of peaks between speaking and listening, $p < 0.0096$, and dotted lines mark peaks with $p<0.05$.

2.3.3 Speaking induced suppression (SIS)

The M100 response in left primary auditory cortex occurred at 92ms and showed SIS ($p = 0.0448$), with a mean SIS magnitude (z-scored), calculated by subtracting the speaking peak value from the listening peak value, equal to 9.264. In right primary auditory cortex, SIS was seen in both the M100 response occurring at 84.8ms (mean magnitude 6.3417, $p = 0.0312$), and in the M200 response occurring at 192.3ms (mean magnitude 2.1731, $p = 0.0184$). Interestingly, the earliest peak (35.6ms) in right primary auditory cortex had anti-SIS ($p = 0.0001$) with a mean magnitude of 6.1625.

Left STG/MTG did not show the SIS effect, but instead showed speaking-induced delays of the M100 response latency, with the M100 occurring 14.7ms later in the speaking condition at 114.5ms than in the listening condition at 99.8ms ($p= 0.0175$). In contrast, in right STG/MTG the m50 showed a listening-induced delay, with the first peak in the speaking condition (37.5ms) preceding the first peak in the listening condition (47.6ms) by 10.1ms ($p= 0.0273$). Right STG/MTG also showed SIS in the m100 response occurring at 87.7ms ($p= 0.0401$), with a mean magnitude of 5.493.

The most prominent and early SIS effects were seen in left vSMG/pSTS, peaking 84.8ms after the speech onset with an average magnitude equal to 5.0239 ($p = 0.0010$). In contrast, right vSMG/pSTS did not have a magnitude difference between the speaking and listening conditions, with the significance values for SIS at the three prominent peaks being $p=0.94$, $p=0.64$, and $p=0.80$. Neither left nor right vSMG/pSTS showed significant latency differences in the peaks between the speaking and listening conditions.

In premotor cortex, one might expect bilateral enhanced activity (i.e., anti-SIS) in the speaking condition compared to the listening condition, especially before speech onset. Yet the only significant enhancement in the speaking condition (magnitude 4.0518, $p = 0.0007$) was

38.3ms after speech onset in the right hemisphere. On the other hand, SIS was seen bilaterally in the premotor cortices. Left premotor cortex showed SIS at 141.2ms (magnitude 2.4004, $p = 0.0357$), while right premotor cortex showed SIS at 74.7ms (magnitude 4.5169, $p = 0.0269$), at 218.9ms (magnitude 1.8163, $p = 0.0459$), and at 292.3ms (magnitude 2.023, $p = 0.0331$).

2.3.4 Cortical neural activity in response to auditory pitch perturbation.

2.3.4.1 Primary auditory cortex responses to auditory pitch perturbation in speaking and listening.

Source time courses for the four bilateral ROI's time-locked to the perturbation onset are shown in figure 2.4. Bilateral primary auditory cortex showed a pronounced three-peak response pattern to the auditory pitch perturbation in both the speaking and listening conditions. This three-peak response, while following the typical MEG auditory response pattern, was delayed as compared to the response pattern to the onset of speech in the first two peaks and in a late response peak. In the listening condition, left primary auditory cortex had peaks at 76.8ms, 140.3ms, 235.2ms following the perturbation onset while in the right hemisphere had peaks at 69.3ms, 148.6ms, 217.6ms following the perturbation onset. Similarly in the speaking condition, left primary auditory cortex had peaks at 96.0ms, 164.8ms, 239.1ms and right primary auditory cortex had peaks at 89.2ms, 146.6ms, and 226.3ms.

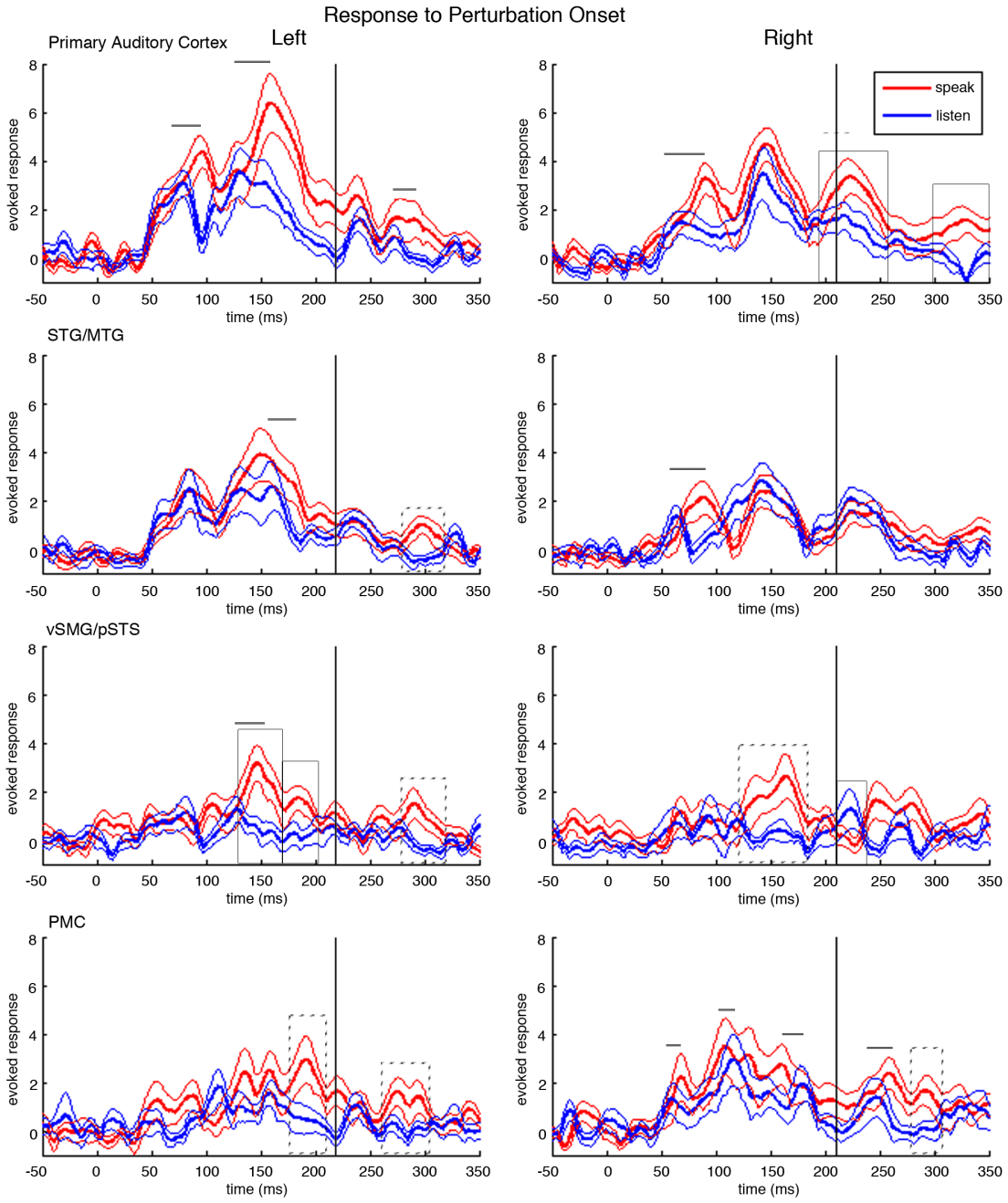

Figure 2.4. Neural responses to the pitch-shifted feedback in both the speaking and the listening condition. In each plot, time is relative to the onset of the pitch-shifted feedback at 0ms. Red traces show the mean cortical responses at the corresponding anatomical location across all subjects in the speaking condition. The thick red line is the grand-average over subjects and the flanking thin red lines are the standard error bars. Blue traces show the mean cortical responses in the listening condition. The thick blue line is the grand-average over subjects and the thin blue lines are the standard error bars. Solid boxes surround peaks with a significant magnitude difference between speaking and listening, $p< 0.0096$, dotted boxes surround peaks with $p<0.05$. Similarly, solid lines mark significantly different latencies of peaks between speaking and listening, $p< 0.0096$, and dotted lines mark peaks with $p<0.05$. The black vertical line denotes the mean time when the subjects' mean pitch time course deviates from the baseline by two standard deviations.

2.3.4.2 Speaking-specific responses prior to and during the onset of compensation.

Responses in bilateral primary auditory cortex showed significant delays in the speaking condition when compared to the listening condition. In the left primary auditory cortex the first peak in the speaking condition (96.0ms) was delayed ($p = 0.0005$) by 19.2ms from the first peak in the listening condition (76.8ms), while in the right primary auditory cortex the first prominent peak in the speaking condition (89.1ms) was delayed ($p = 0.00001$) by 19.8ms from the peak in the listening condition (69.3ms). In the second peak, left primary auditory cortex showed a 24.5ms delay in the speaking condition (164.8ms) to the listening condition (140.3ms), $p = 0.0002$.

Bilateral STG/MTG had speaking induced delays, but at different latencies. The speaking condition in left STG/MTG had a peak at 178.2ms, 13.3ms delayed to the corresponding peak in the listening condition at 164.9ms ($p = 0.0027$). In contrast, right STG/MTG showed speaking induced delays early, with the first peak in the speaking condition occurring at 85.3ms, 21.2ms delayed to the peak in the listening condition at 64.1ms ($p = 0.0002$).

Left vSMG/pSTS had three distinct peaks with significant SPRE. The first, and largest SPRE in left vSMG/pSTS occurred at 149.6ms with a magnitude of 2.2757 ($p = 0.0033$). This SPRE had a corresponding speaking induced delay, with the peak in the speaking condition (149.6ms) 22.9ms after the peak in the listening condition (126.7ms), $p < 0.0001$. The following peak, at 186.7ms, also showed significant SPRE with a magnitude of 2.2304 ($p = 0.005$). The third SPRE peak in left vSMG/pSTS occurred following the compensation. Right vSMG/pSTS showed one wide SPRE peak at 167.3ms, with a magnitude of 2.3852 ($p = 0.0306$).

Bilateral premotor cortices (PMC) showed increasing activity in response to the pitch shifted feedback in both the speaking and the listening condition after the perturbation onset. In

left premotor cortex the response in the speaking condition had a greater magnitude and duration than the response in the listening condition. Left premotor cortex showed a SPRE response at 190.1ms with a magnitude of 2.9751 ($p = 0.0399$). Right premotor cortex had several peaks with significant speaking-induced delays and a late peak shows SPRE. The first peak with a speaking induced delay occurred at 66.8ms in the speaking condition, 8.4ms after the peak in the speaking condition at 58.4ms ($p = 0.004$). Later, the peak in the speaking condition at 125.1ms followed the corresponding peak in the listening condition (109.9ms) by 15.2ms ($p = 0.002$). Finally, the peak in the speaking condition at 181.9ms had a corresponding peak 21.2ms earlier in the listening condition at 160.7ms ($p < 0.0001$).

2.3.4.3 Speaking-specific responses following the onset of compensation.

Following the compensation, left primary auditory cortex showed one speaking specific response, a 16.6ms speaking-induced delay ($p = 0.0001$), with the peak in the speaking condition occurring at 288.1ms and the peak in the listening condition occurring at 271.5ms. Right primary auditory cortex showed a 8.7ms speaking-induced delay between the peak in the speaking condition at 226.3ms and the peak in the listening condition at 217.6ms ($p = 0.0120$). SPRE was observed in right primary auditory cortex at 226.3ms after the perturbation with a magnitude of 1.583 ($p = 0.0062$), and at 326.2 ms (magnitude= 2.4026, $p = 0.0064$).

The only speaking specific response in bilateral STG/MTG following the compensation occurred in left STG/MTG. Left STG/MTG showed SPRE in a peak 300.4ms after the perturbation onset with a magnitude of 1.9203. Similarly, the only SPRE peak in vSMG/pSTS following the compensation was in left vSMG/pSTS. A late peak, occurring at 291.6ms, showed SPRE with a magnitude of 2.1761 ($p = 0.0019$). Conversely, right vSMG/pSTS had greater

magnitude in the listening condition as compared to the speaking condition at 220.5ms with a magnitude of 1.9119 ($p = 0.0456$).

Left premotor cortex also showed SPRE in a late peak at 276.7ms, with a magnitude of 1.9772 ($p = 0.0229$). Similarly, right premotor cortex showed SPRE in the peak at 288.9ms showed SPRE (magnitude = 1.7373, $p = 0.023$). A late peak in right premotor cortex during the speaking condition (258.2ms) preceded the corresponding peak in the listening condition (243.5ms) by 14.7ms ($p< 0.0001$).

2.3.5 Correlation of cortical activity with behavioral compensation.

Correlations between an individual subject's behavioral compensation and cortical activity during speaking for each peak showing SPRE were calculated. One SPRE peak showed significant correlation across subject compensation, the peak in right premotor cortex with the mean latency of 288.9ms. This response in this cortical area showed a significant, yet small, negative linear regression with behavioral compensation ($p=0.049$, r-squared = 0.401). The peak at 186.7ms in left vSMG/pSTS conversely showed a trend towards a significant positive correlation with compensation ($p=0.098$, r-squared = 0.305).

2.4 Discussion

In this study, we used MEG to examine auditory feedback processing within a subset of the speech sensorimotor network. This study addresses an integral part of speech, phonation. Phonation conveys important prosodic and affective information that are an invaluable part of speech. Therefore understanding the neural circuitry that controls phonation is important and relevant to the understanding of speech. In this study, we looked at responses to speech onset and to brief, unexpected perturbations of pitch feedback, and compared responses seen during

speaking with those seen during subsequent listening to playback of the feedback. In our analysis of speech onset responses, we found that speaking-induced suppression (SIS) occurs well beyond auditory cortices. While we replicate SIS in bilateral primary auditory cortex, we found that the most significant SIS effect occurs in left vSMG/pSTS, and interestingly, we document for the first time speaking induced delays in left STG/MTG. In our analysis of perturbation responses, we found speech perturbation response enhancement (SPRE) in a large bilateral network of regions, with maximal SPRE seen in the left vSMG/pSTS but also including right primary auditory, left STG/MTG, right vSMG/pSTS, and bilateral premotor cortex. We also observed significant speaking induced delays in perturbation responses within bilateral primary and secondary auditory cortices, left vSMG/pSTS, and right premotor cortex. We found little evidence of a correlation between SPRE and compensation across subjects. We discuss these findings in the context of the growing literature on responses to unaltered and altered speech feedback.

2.4.1 The effect of the onset of speech on MEG magnitude response.

While previous work studying the neural response to unaltered self-produced speech has demonstrated SIS in auditory areas, these previous studies did not examine the spatial distribution of the SIS response beyond auditory cortices (Curio et al., 2000; Houde et al., 2002; Ventura et al., 2009). In this study we explored four nodes of the speech sensorimotor pathways and found SIS responses at a number of locations. Although SIS responses were seen in bilateral primary auditory cortex and right STG/MTG, the most prominent SIS responses were localized primarily to left vSMG/pSTS. Previous reports averaged sensor-space data from multiple sources across temporal cortex (Curio et al., 2000; Houde et al., 2002), and so include responses from sources in both primary auditory cortex and vSMG/pSTS. These studies also found weaker

SIS effects in the right hemisphere. In this study, the SIS in the responses across the right temporal cortex did not survive conservative correction for multiple comparisons. Yet, if these responses from right primary and secondary auditory cortices were summed, significant SIS could occur. Of interest is the large early response (35.6ms) in right primary auditory cortex in the speaking condition. This temporally and spatially specific anti-SIS response combines with the surrounding SIS to create a pattern consistent with the distributed representation of SIS reported by Greenlee and colleagues using electrocorticography (ECoG) (Greenlee et al., 2011).

2.4.2 The effect of speaking on MEG response latency.

Speaking specific responses to the onset of speech were not limited to SIS. Significant speaking induced delays were also found in this study. Even though latency differences have not received much attention in the speech production literature, their existence here suggests the possibility of additional speaking-specific processing, where information is conveyed by MEG peak delays in the speaking condition. Indeed, prior work in speech perception has already shown that peak latency can carry important information. For example, latency of the m100 response during passive listening varies significantly across vowels or across consonants (Gage et al., 1998; Poeppel et al., 1997). In this study, a speaking-induced delay was found in secondary auditory areas (STG/MTG) in response to the onset of unaltered feedback. Previous studies have reported conflicting information on a latency difference between responses in active speaking and passive listening conditions. Curio et al. report that the m100 in the speaking condition was significantly delayed compared to the m100 in the listening condition in bilateral auditory areas (Curio et al., 2000). In another study, Houde et al. did not report a latency difference between the speaking and the listening conditions with unaltered auditory feedback, but when feedback

was altered to produce noisy auditory feedback, the speaking condition was delayed (Houde et al., 2002).

Speaking induced delays could also play an important role in the preliminary auditory processing of the pitch-altered feedback. The results in this study show discrete delayed responses to the altered feedback in the speaking condition when compared to the peak latencies in the listening condition. Given that a change in fundamental frequency is a weaker acoustic event than the onset of speech, it is not surprising that the three-peak auditory response to the perturbation in both the speaking and listening conditions were delayed as compared to the three-peak response to the onset of speech. Yet, of particular interest are the peaks that show an additional selective delay in the speaking condition in response to the altered auditory feedback. These significant delays demonstrate processing that differs both from passive listening to the pitch shift and from speaking with unaltered feedback. These speaking induced delays in response to the altered auditory feedback were seen in the early responses in bilateral primary auditory cortex, bilateral STG/MTG, left vSMG/pSTS and in several peaks in right premotor cortex. The first speaking specific responses to the perturbation are the speaking induced delays in right primary auditory cortex and right STG/MTG followed shortly by left primary auditory cortex. The right temporal cortex has previously been demonstrated to be important in pitch perception (Samson and Zatorre, 1988; Sidtis and Volpe, 1988; Zatorre and Halpern, 1993). Importantly, the speaking induced delays cannot be attributed simply to a change in pitch, since previous studies have shown that varying the fundamental frequency of auditory stimulus is not sufficient to alter the latency of the m100 (Poeppel et al., 1997). One intriguing possible consequence of the observed speaking-induced delay in right primary auditory cortex and STG/MTG is additional processing in the speaking condition involving the recognition and initial processing of the pitch feedback error.

The speaking induced delays we observed might also help to explain why some results, such as the 2009 study of Behroozmand et al. differ from our results. In their previous EEG study, Behroozmand and colleagues found enhanced evoked responses in bilateral scalp electrodes to pitch-shifted auditory feedback during active vocalization compared to passive listening in both the P1 (73.51ms) and P2 components (199.55ms) (Behroozmand et al., 2011a). While Behroozmand et al. do not report behavioral compensation; both SPRE peaks precede the onset of compensation found in our study. In the current study we do not see the early enhancement (73.51ms) previously reported, but we do see significant speaking induced delays in the early response components. These speaking induced delays, when summed over cortical regions and recorded on the scalp, could appear as a magnitude difference.

2.4.3 Altered auditory feedback induces behavioral compensation.

The behavioral effect of the unexpected pitch shift during auditory feedback induces a rapid compensatory response that, on average, opposes the direction of the external shift for each subject. The mean compensation varied across subjects, replicating findings in previous studies (Burnett et al., 1998; 1997; Greenlee et al., 2013). The response latency measured in our study, with a mean of 210.9ms, is the same as one previous study, and slightly later than some previously reported mean latencies, but within the range of latencies. Previous studies have reported mean latencies of 210.91ms, standard error 31.84 (Jones and Munhall, 2002), 159ms, range 104-223ms (Burnett et al., 1997), 192ms (Burnett et al., 1998), and 184ms, range 115ms-285ms (Greenlee et al., 2013) and as early as around 100ms (Korzyukov et al., 2012; Larson et al., 2007). The point where the F0 contour crosses the two standard deviation line is a conservative measure of when the compensation has definitively begun, but is likely later than the actual initiation of the new vocalization. Thus, in order to account for this discrepancy, neural

responses immediately preceding the compensation onset will be considered here as occurring during the onset of compensation.

Some previous work has shown that long duration pitch shifts can induce a multipeak response (Burnett et al., 1998; Hain et al., 2000). The two phase response was most clearly reported in a previous study in which subjects were specifically instructed to make a voluntary response to a pitch perturbation (Hain et al., 2000). We did not see a two-component response, which may be due to the fundamental task difference, since our subjects were neither instructed to respond to nor ignore the pitch alteration, and due to the fact that we looked at all trials averaged for each subject.

2.4.4 The effect of altered auditory feedback in speech on MEG magnitude response.

In addition to the speaking induced delays, speaking specific neural enhancement occurred in later cortical responses. As opposed to other studies that found an unlocalized SPRE response in two early peaks (Behroozmand et al., 2011b; 2011a; 2009; Korzyukov et al., 2012), we found maximal SPRE responses localized to left vSMG/pSTS, and additional SPRE distributed throughout the bilateral speech motor network. The sequence of SPRE can be divided into two intervals, the SPRE occurring before and during the behavioral compensation at 210.9ms and the SPRE occurring after onset of the behavioral compensation. We discuss these intervals separately because models of speech motor control postulate different processes are occurring during these different intervals (Hickok et al., 2011; Houde and Nagarajan, 2011; Ventura et al., 2009). According to such models, the SPRE occurring before and during the compensation is likely to be involved in the recognition of the speech error, the processing of that error and the motor plan change that causes the subsequent behavioral compensation. The SPRE occurring after the onset of the compensation is likely to be involved in the continuously

changing motor command as the compensation increases, preparing for the new auditory input of the compensated pitch, and monitoring the new auditory feedback due to the updated motor plan.

2.4.4.1 Role of SPRE Responses: Prior to and during compensation initiation.

Two significant SPRE peaks occur prior to and during the initiation of compensation in left vSMG/pSTS and one SPRE peak occurs in right vSMG/pSTS, a region encompassing areas of parietal cortex along the SMG and area Spt (Sylvian fissure at the parietal-temporal boundary). These areas have been implicated as crucial in sensorimotor translation (Andersen et al., 1997; Hickok et al., 2011; 2009; Shum et al., 2011). Given the previously established role of vSMG/pSTS in auditory motor translation, we interpret this enhancement as encompassing processing related not only to recognizing the pitch error, but also crucial computation to both induce and prepare for the resulting motor response. Further support for the role of vSMG/pSTS in changing the motor command is the SPRE immediately following in left premotor cortex. The SPRE in premotor cortex would be most naturally implicated in driving the compensatory motor response, but it is important to ensure that this assumption is causally reasonable given the timing of the SPRE compared to the compensation. Rödel et al found activation of the vocalis muscle 10.7-11.8 ms following cortical stimulation and activation of the cricothyroid muscle 10.1-10.8ms following cortical stimulation (Rödel et al., 2004). Given that work, the timing of SPRE in the premotor cortex reasonably allows for SPRE to be involved in compensation.

2.4.4.2 Role of SPRE Responses: Following compensation onset.

The first SPRE response following compensation occurs in right primary auditory cortex. It is probable that this speaking specific activity is once again demonstrating the well-established role of the right hemisphere in pitch detection. Interestingly, following this pitch detection we observe a series of SPRE originating in left premotor cortex followed by right premotor cortex, and following the speech motor control network via left vSMG/pSTS through left STG/MTG and back to right primary auditory cortex. While a causal link cannot be ascertained with these data, the temporal sequence suggests efference copy of a corrected motor plan being fed back to auditory cortex. Consistent with this interpretation, we note that while the compensation onset occurs at 210.9ms, the compensation continues to ramp up for the duration of the pitch shift until the peak at 531.8ms.

2.4.4.3 Correlation of MEG response during speaking with behavioral compensation.

Interestingly, our results did not show a strong positive correlation between cortical response to the perturbation in the speaking condition and individual subject compensation at the peaks that exhibit SPRE. While we did see a trend towards a positive correlation in left vSMG/pSTS, we did expect a stronger association within this region. Furthermore, we expected to see a positive correlation in premotor cortex with compensation. Instead, activity in left premotor cortex did not show any relationship with compensation, while right premotor cortex showed a weak negative correlation with compensation. These findings suggest that information influencing the degree of compensation was not captured in the low-frequency evoked response data analyzed in this study. In contrast, Chang et al. did find cortical areas whose activity correlated with compensation (Chang et al., 2013). What accounts for this difference in findings? First, Chang et al. looked for correlations with activation in the high gamma (50-150Hz)

frequency band, instead of the lower frequency responses examined here. Chang et al. were also analyzing trial-to-trial correlations within individual subjects, whereas here we correlated mean responses with compensation across subjects. Finally, in the Chang et al. study, the correlation analysis was response locked, with the activity time course of each trial being aligned to the time of peak compensation for that trial, whereas here trials were aligned to perturbation onset before being averaged. Overall, the Chang et al. study implicates high gamma responses are important in driving single trial variability in compensation within subjects. Whether mean compensation is also correlated to mean high gamma responses across subjects is yet to be addressed.

2.4.5 Relationship between SIS and SPRE

In addition to looking at the temporal sequence of SPRE, the design of this study also permitted study of the relationship of SIS and SPRE. The SIS response showed dissociation from the SPRE response in several cortical areas. Left STG/MTG and right vSMG/pSTS showed SPRE, but did not have a SIS response. Left primary auditory cortex and right STG/MTG showed SIS, but did not show SPRE. Right primary auditory cortex and bilateral premotor cortex showed both SIS and SPRE, but these peaks did not survive corrections for multiple comparisons. Only one region showed co-localization of a significant SIS and SPRE response, left vSMG/pSTS, the sensorimotor translation region.

The first attempt to relate cortical responses to unaltered vocalizations with responses to a feedback error was conducted in work done in non-human primates. Eliades and Wang found that the majority of neurons in the auditory cortex of marmoset monkeys (approximately three-quarters) were suppressed during vocalization, compared with their pre-vocalization activity (Eliades and Wang, 2008). Overall, the neurons that showed this suppression during vocalization showed increased activity during frequency-shifted feedback but maintained suppressed activity

as compared to the period preceding the vocalization. These findings were interpreted to say that the vocalization induced suppression increases the neurons sensitivity to vocal feedback. As stated above, we did not see this direct relationship between SIS and SPRE in human left auditory cortex but we did see both SIS and SPRE in right primary auditory cortex. Instead, SIS and SPRE co-localized in left vSMG/pSTS. Fundamental differences in the task and definition of suppression and enhancement can account for the differences in results between these studies. While in the current study suppression and enhancement in speaking were defined in comparison to the neural response to a listening condition with identical auditory input, in the Eliades and Wang study suppression and enhancement were defined as the comparison between the active window and the pre-vocalization activity. The current study finds the population response in left vSMG/pSTS to show both SIS and SPRE, but further studies are necessary to test if the suppression at the onset of speech increases the neurons' sensitivity to a feedback error.

One previous study has looked at the spatial relationship with SIS and SPRE in ECoG in the left hemisphere (Chang et al., 2013). This study found the majority of electrodes preferentially demonstrate SIS or SPRE, with a few overlapping electrodes scattered across left temporal cortex, left inferior parietal lobe, and one electrode in left ventral premotor cortex. The majority of the electrodes showing overlapping SIS and SPRE are clustered near the temporal parietal junction, a region encompassed in the vSMG/pSTS region in the current study, where SIS and SPRE were colocalized. This overlapping, yet dissociated pattern of responses to the unaltered speech onset and the pitch perturbation shows that the networks involved in recognizing self-produced speech are similar to, but not identical to, those involved in recognizing and responding to an error in auditory feedback. For several models of speech motor control, these results are somewhat problematic, since such models assume the processes responsible for SPRE get their input from processes responsible for SIS (Chang et al., 2013). Thus, these models

predict all cortical areas showing SPRE should also show SIS, which the results here show is not always the case.

2.4.6 Limitations

In order to understand speaking-specific cortical activity, we use passive listening to the subject's recorded voice from the previous block as a reference signal with which to discuss responses during speech. The auditory input during the passive listening matches the auditory input through the headphones during speech, but does not include the additional sound perceived during speaking by bone conduction. The minimal role of bone conduction in SIS has been indirectly addressed by a study demonstrating that motor induced suppression occurs in response to auditory input created by a volitional motor act (Aliu et al., 2009). In that study, where the auditory input was perfectly matched in the active and passive listening conditions, subjects experienced motor induced suppression when hearing a tone in response to a button press. It is thus unlikely that the suppression seen during speech is an artifact of bone conduction instead of speech motor induced suppression. The impact of bone conduction on the pitch perturbation response has also been well addressed by Burnett and colleagues in their 1998 study (Burnett et al., 1998). That study found the addition of pink masking noise to block the effect of bone conduction did not change the pitch perturbation response, importantly showing that bone conduction does not play a role in the perturbation response. The current study was designed to examine the role of auditory feedback in speech production, so a closely matched auditory task works as an effective reference signal. Yet during ongoing speech there are also complex and dynamically changing somatosensory cues influencing speech. The interplay between the role of somatosensory feedback and auditory feedback during speech production, which has only

recently been addressed in a behavioral study (Lametti et al., 2012), is an exciting question for future studies.

2.5 Conclusions

Results from the present study suggest that there are distinct, yet overlapping networks involved in monitoring the onset of speech and in detecting and correcting for an error during ongoing speech. SIS was found in bilateral auditory cortex and left ventral supramarginal gyrus/posterior superior temporal sulcus (vSMG/pSTS). In contrast, during pitch perturbations, activity was enhanced in bilateral vSMG/pSTS, bilateral premotor cortex, right primary auditory cortex, and left higher order auditory cortex. The results in the study also suggest that the latency of cortical responses may also be important in understand cortical processing during speech. Importantly, vSMG/pSTS is involved in both monitoring the onset of speech and recognizing and responding to auditory errors during speech.

Chapter 3.

Neural responses during speech production at vocalization onset demonstrate cortical self-monitoring.

3.1 Introduction

One characteristic of cortical monitoring of one's own auditory feedback is the suppression of cortical regions to self-produced vocalizations at their onset. This cortical suppression at the onset of vocalization, generally termed Motor Induced Suppression (MIS), and more specifically Speaking Induced Suppression (SIS), has been demonstrated in non-human primates with single-unit recording (Eliades and Wang, 2005; Muller-Preuss and Ploog, 1981) and in humans with intracranial recording (ECoG) (Chang et al., 2013; Flinker et al., 2010; Greenlee et al., 2011), electroencephalography (EEG) (Behroozmand and Larson, 2011), in magnetoencephalography (MEG) (Curio et al., 2000; Houde et al., 2002; Kort et al., 2014; Niziolek et al., 2013) and in fMRI (Agnew et al., 2013). Evidence that SIS is a measure of auditory self-monitoring comes from the finding that SIS is reduced when the true auditory feedback does not match the expected auditory feedback (Behroozmand and Larson, 2011; Heinks-Maldonado et al., 2006). These studies did not employ source localization so the results average responses originating throughout and beyond the temporal lobe. Therefore, the question remains whether the reduction of SIS with altered auditory feedback is either 1) a true reduction of suppression by the increased activity during speaking or is instead 2) the superposition of suppression and enhancement.

The anatomy of SIS monitoring is disputed, and has been reported in the temporal lobe anterior to Heschl's gyrus and inferior parietal lobe (Agnew et al., 2013), primary auditory cortex and supramarginal gyrus (Kort et al., 2014), across the temporal lobe (Greenlee et al., 2011) and

into inferior parietal cortex (Chang et al., 2013). To add further complications, within the same subject, the location of SIS varies according to whether it is measured by looking at an evoked, low frequency analysis or an induced, high gamma band analysis (Greenlee et al., 2011).

In this study, we used magnetoencephalography imaging (MEG) to examine the spatial and frequency distribution of SIS on the entire cortex. By using excellent source localization techniques we were able to test if individual regions of cortex show reduction of SIS when the auditory feedback deviates both slightly and considerably from the expected acoustic consequence of speaking. Further, we explored whether the reduction of SIS is driven by increased activity during speaking with pitch altered auditory feedback or a superposition of suppression and enhancement across the temporal lobe.

3.2 Materials and Methods

3.2.1 Subjects

Twelve right-handed (6 female) English speaking volunteers with normal speech and hearing participated in this study. All participants gave informed consent after procedures had been fully explained. The study was performed with the approval of the University of California, San Francisco Committee for Human Research.

3.2.2 MEG Recording

The task was completed during whole head MEG neural recording in awake subjects laying in the supine position. The MEG system (MISL, Coquitlam, British Columbia, Canada) consists of 275 axial gradiometers and was recorded with a sampling rate of 1200Hz. Three fiducial coils were placed on the nasion and left/right preauricular points to triangulate the position of the head relative to the MEG sensor array. In a separate session high resolution

anatomical MRIs were obtained for each subject. The fiducial markers points were later co-registered with an anatomical MRI to generate head shape.

3.2.3 Experimental Design and Procedure

The experiment was administered in a block design with a Speaking Condition and Listening Condition. In both the Speaking Condition and the Listening Condition, subjects watched a screen with a projected image in their line of sight. The background of the screen was black. A trial was initiated when the words "say 'ah'" appeared on the screen.

In the Speaking Condition, subjects were instructed to produce a brief utterance of the vowel /a/ by the experimenter before the initiation of the block. The subjects spoke into an MEG-compatible optical microphone and received auditory feedback through insert earphones. During the first 75 of 450 trials subjects received unaltered auditory feedback. For the remaining 375 trials, feedback pseudorandomly alternated between unaltered feedback (0cent), +/-100cent, and +/-300cent alteration. In the Listening Condition, subjects were instructed to observe the screen but to passively listen despite receiving the same visual prompts. Subjects passively listened to the recording of their perturbed voice feedback obtained in the previous Speaking Condition block. The auditory input through earphones in both conditions was identical. Prior to the start of the experiment, the volume of auditory input through the earphones was adjusted so that subjects reported their auditory feedback was the same as or slightly louder than expected.

3.2.4 Audio analysis

The pitch analysis was performed on each subject's single trial audio data. The microphone and feedback signals were recorded and analyzed, sampled at 11025Hz. The data was recorded in 32-sample frames. Pitch was estimated for each of these frames using the

standard autocorrelation method (Parsons, 1986). The resulting frame-by-frame pitch contour was then smoothed with a 20Hz, 5th order, low pass Butterworth filter. Using visual inspection, trials with erroneous pitch contours were removed. The grand average across subjects was calculated by detrending an individual subject's mean pitch and then averaging across subjects. The first 10% and final 5% of each subject's trials were removed to eliminate edge effects causing errors in the pitch tracking. Only the interleaved 0 cent trials were used to for comparison to allow for equal numbers of trials for each condition.

3.2.5 MEG data preprocessing

The MEG sensor data were marked at the speech onset and at the perturbation onset. Third gradient noise correction filters were applied to the data and the data were corrected for a DC offset based on the whole trial. Artifact rejection of abnormally large signals due to EMG, head movement, eye blinks or saccades was performed qualitatively through visual inspection and trials with artifacts were eliminated from the analysis. Sensor data was notch filtered around 120Hz with a width of 4Hz.

3.2.6 MEG data analysis

3.2.6.1 Sensor analysis.

MEG sensor analysis was done by averaging across all trials, and then over either left or right temporal sensors. Results were filtered 2-40Hz and the root-mean-square of the data was calculated. Each subject's data was normalized using 100ms prestimulus period to create a z-score. During the prestimulus period subjects were observing a blank screen between trials. A grand-average was then calculated across subjects. Peak activity for each subject was calculated by finding the peak in the grand-averaged listening data and then finding the corresponding peak

in the individual subject's data within a 40ms window centered at the grand-average peak. T-tests were then performed between speaking condition and listening condition and values $p<0.05$ were reported.

3.2.6.2 Source-space Time-Frequency Analysis.

The NUTMEG time-frequency spatially adaptive filter algorithm (Dalal et al., 2008) was used to localize the induced activity in the theta/alpha (4-13Hz), beta (13-30Hz), gamma (30-55Hz) and high gamma band (65-150Hz) to the individual subject's spatially normalized MRIs. The NUTMEG algorithm has been previously described in detail in Dalal et al, 2011 (Dalal et al., 2011). A brief description of this method follows. The multisphere lead field (forward model) was calculated for every 5mm voxel in the brain. The lead field describes the magnetic field strength at each MEG sensor that would arise from a single dipole source originating in each voxel. Source localization was then calculated for the high gamma band activity using both the lead field and sensor covariance. Noise corrected pseudo-F ratios were computed between the active windows and the prestimulus control baseline. Voxels in deeper brain structures were removed, restricting further analyses to the cortical surface.

For theta/alpha band, one 400ms window was analyzed. For beta and gamma bands, one 200ms time window was analyzed. For high gamma band, two 100ms time windows were analyzed- 0-100ms and 100-200ms following speech onset. Lower frequency bands require larger analysis windows to have sufficient information to extract band-specific activity. Only single time windows were analyzed for the lower frequencies to restrict the analysis to the window during which the subjects were vocalizing.

Group statistics were computing using the NUTMEG time-frequency statistics toolbox (Dalal et al., 2011; 2008) with statistical non-parametric mapping (SnPM) (Singh et al., 2003).

The results from the time-frequency beamformer for each subject were normalized to the MNI template brain using SPM2. For every time-frequency point, three-dimensional average and variance maps were calculated and then smoothed with a 20x20x20 mm^3 Gaussian kernel. Using this, a pseudo-t statistic was obtained at each voxel, time window, and frequency band. Voxel labels were permuted 2*Number of subjects times to create nonparametric null distributions to derive p-values. The neuro-behavioral correlations were calculated by computing Pearsons' correlation coefficients for activations for all voxels again behavioral compensation. A cluster correction (20 voxel cluster, $p<0.05$) was then applied to correct for multiple comparisons. For the speaking vs. listening conditions comparison in the high gamma band, a Bonferroni correction was applied for more stringent statistical thresholding. In the across condition comparisons, analyses were masked to only include regions that showed a significant change from baseline in either the speaking condition or listening condition.

3.3 Results

3.3.1 Brief ballistic utterances do not show compensation.

Subjects produced brief, ballistic utterances of the vowel /a/ in response to the visual prompt. Utterances had a mean duration of 314ms with a standard deviation of 96.7ms. Subjects did not compensate for the altered auditory feedback. The grand average of pitch production across subjects and an example subject are shown in figure 3.1 showing no difference in produced pitch in shifted trials compared to unshifted trials.

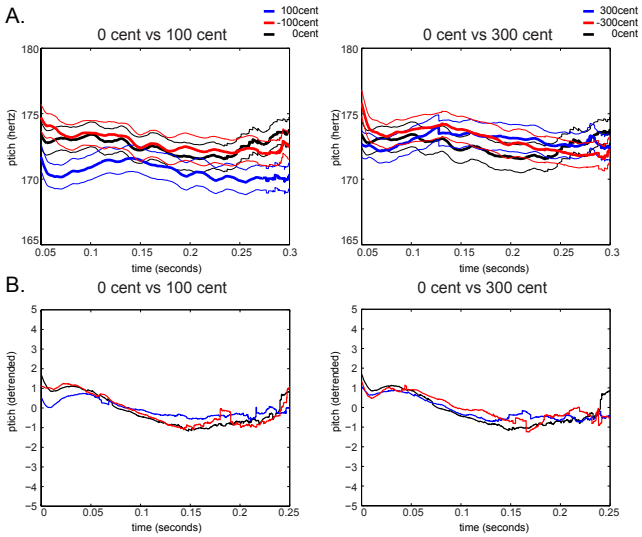

Figure 3.1. Vocal responses to pitch-altered auditory feedback do not show compensatory changes. A. Top panel shows mean (thick lines) with standard error bars (thin lines) of a single subject's pitch production. The subject's pitch production during unaltered auditory feedback is shown in black, during altered feedback that increases the subject's pitch in blue, during altered feedback that decreases the subject's pitch in red. Pitch productions during alterations 100 cent in magnitude are shown in the left column; pitch productions during alterations 300 cent in magnitude are shown in the right column. B. The grand-average of the pitch production (linear trends removed) across all subjects.

3.3.2 Evoked sensor responses reveal SIS in bilateral temporal regions.

Evoked responses are discrete neural events time locked to speech onset. Evoked responses averaged across left and right temporal sensors showed SIS when speaking with unaltered feedback, figure 3.2. In the left hemisphere temporal sensors the standard three peak auditory response was present in the listening condition with unaltered feedback with peaks at 50.8ms, 93.3ms and 155.0ms. In the speaking condition with unaltered feedback the corresponding peaks were significantly reduced compared to listening, indicating SIS in all three peaks. In the right hemisphere SIS occurred in the broad m100 peak (with two peaks, one at 80.0ms and one at 100.8ms) and in a later peaks at 121.7ms and 198.3ms. When the subjects' auditory feedback was altered by either raising or lowering the pitch by 100 cents, SIS was observed in the left temporal sensors at the 53.3ms, 95.8ms, 102.5ms and 185.0ms peaks, but was

absent in the right hemisphere in all but the single late peak at 187.5ms. When the subjects' auditory feedback was altered by 300 cents, a large shift in pitch, the SIS was eliminated bilaterally with the exception of one late peak in the left temporal sensors at 190.8ms.

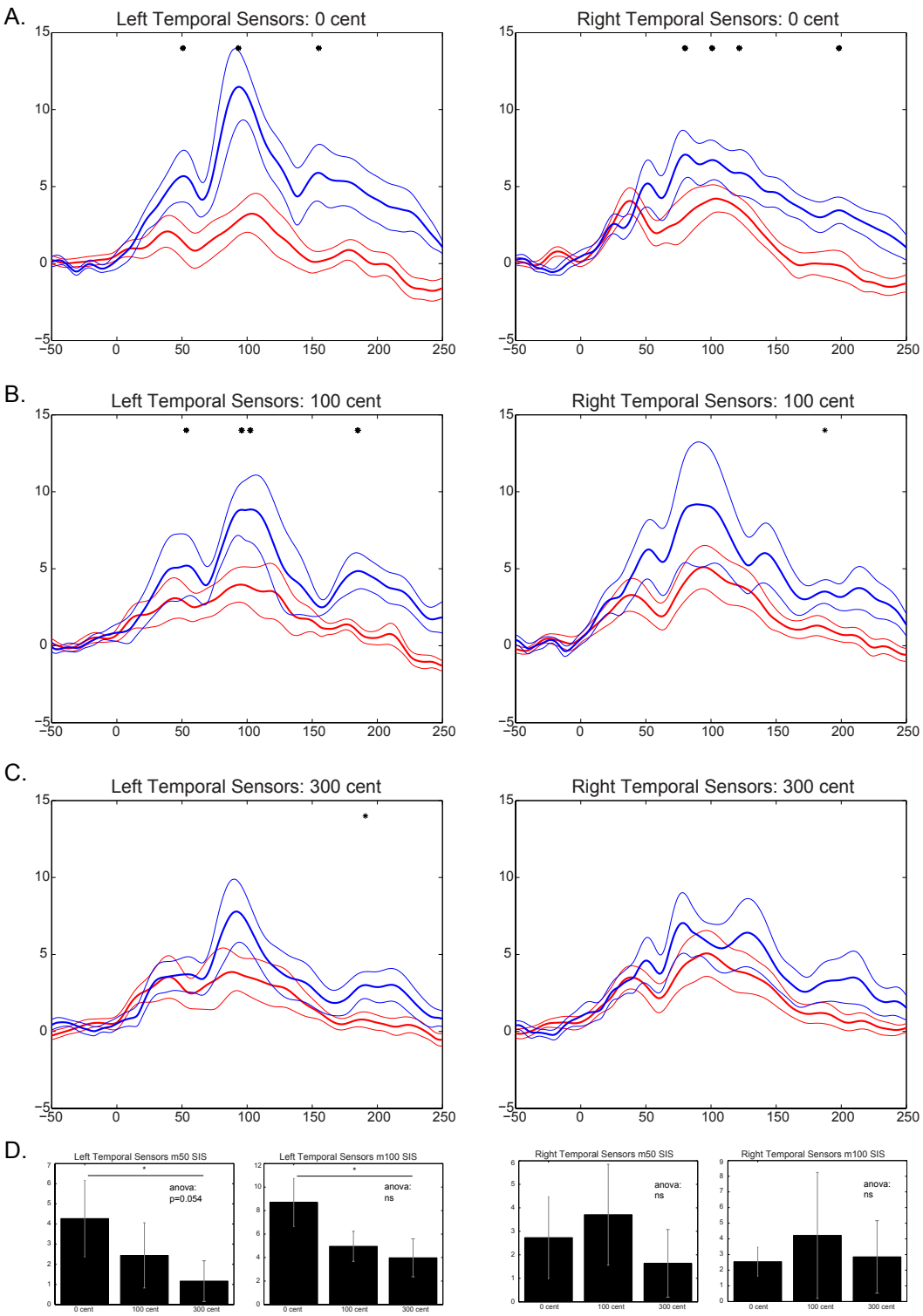

Figure 3.2. Average low-frequency evoked responses averaged over left temporal MEG sensors (left column) and right temporal MEG sensors (right column). A. Temporal evoked responses during speaking with unaltered auditory feedback shows SIS. B. Temporal evoked responses during speaking with 100 cent altered auditory feedback shows reduced SIS. C. Temporal evoked responses during speaking with 300 cent altered auditory feedback shows the absence of SIS. D. Bar graphs show the magnitude of SIS in the left temporal sensors is modulated with the amount the auditory feedback deviates from the expected auditory feedback. Stars denote peaks with $p < 0.05$.

3.3.3 Time-frequency analysis shows location and frequency band specific patterns of SIS.

Previous work into SIS has primarily used evoked, broadband low frequency SIS, and has not addressed how cortical activity is modulated across frequency bands in the whole brain. Not surprisingly, when cortical activity in the whole brain is analyzed a very complex dynamic of SIS is present not just across cortical regions, but even across frequency bands, figure 3.3. With MEG induced analysis, both power increase and decrease can highlight important task related cortical processing (Pfurtscheller and Lopes da Silva, 1999). Since both power increase and decrease are important, we reported induced SIS as occurring when the absolute value of the cortical response in the speaking condition is significantly less than the absolute value of the response in the listening condition. Figure 3.3 shows significant differences between the speaking condition and the listening condition with unaltered (0 cent) feedback, 100 cent altered feedback and 300 cent altered feedback with suppression in the speaking condition (SIS) displayed in cool colors (blue to green) and enhancement in the speaking condition (anti-SIS) displayed in warm colors (red to yellow). Table 3.1 lists the corresponding peak voxels from the comparison of speaking and listening with unaltered auditory feedback, and whether the SIS was present with in the conditions with altered auditory feedback. It is to be expected that many regions involved in speech production, including motor, premotor and somatosensory regions would show greater activity during speaking than during listening, so we restrict our discussion to regions that show significant SIS.

In a broad low frequency band encompassing both alpha and theta only two right hemisphere regions showed SIS, a cluster in the right posterior temporal lobe across both middle and inferior temporal gyri and right cerebellum. Both of these SIS clusters were absent with altered auditory feedback.

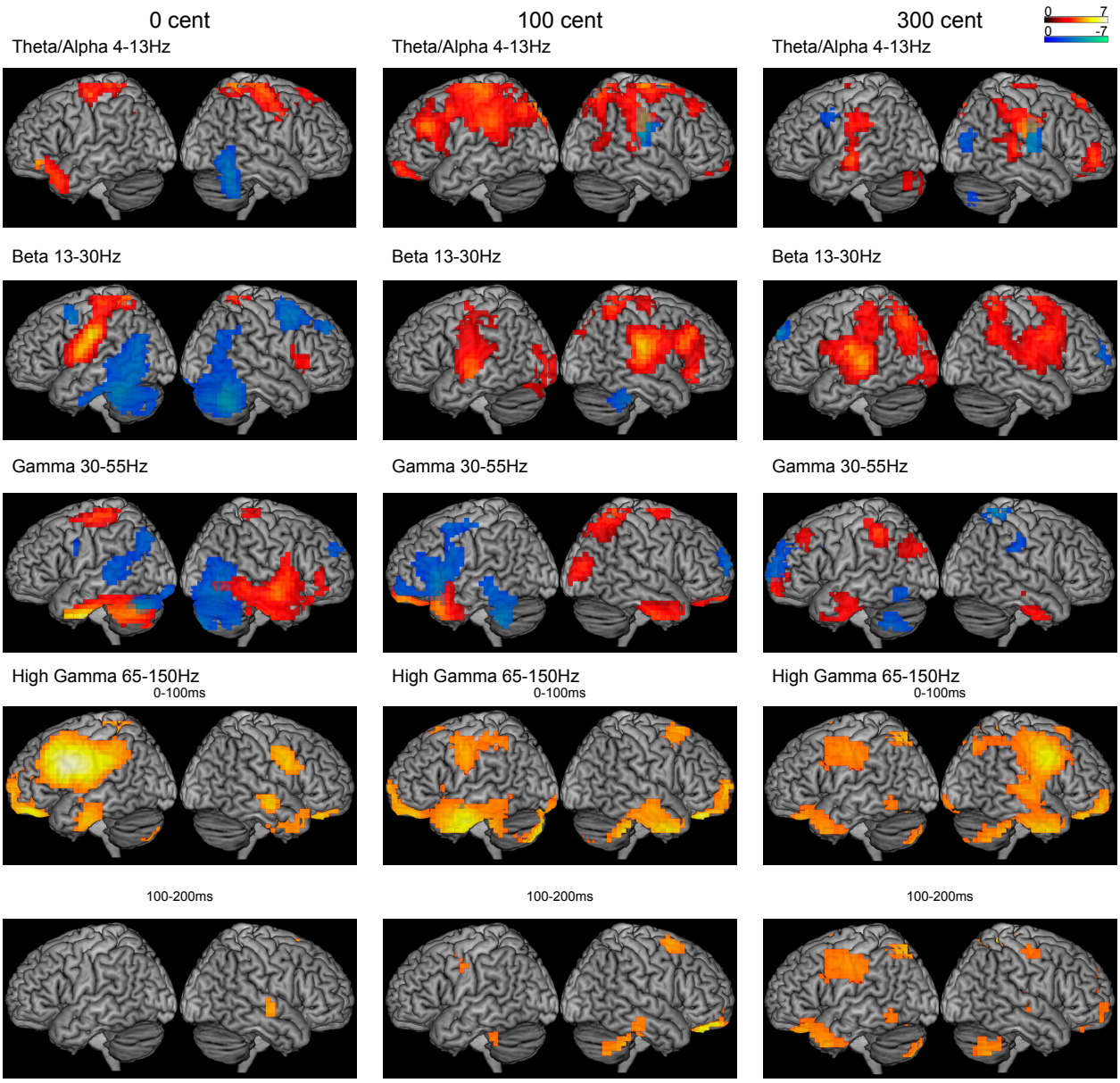

Figure 3.3. Cortical induced activity that significantly differed between the speaking condition and listening condition across frequency bands. Left column shows responses to speaking and listening with unaltered auditory feedback, middle column shows cortical responses to speaking and listening with 100 cent altered auditory feedback, left column shows cortical responses to speaking and listening with 300 cent altered auditory feedback. Cool colors, blue to green, show cortical regions that were suppressed during speaking compared to passive listening (SIS) while warm colors, red to yellow, show enhanced activity during speaking compared to passive listening (anti-SIS). Figures corresponding to frequency bands theta/alpha, beta, and gamma were cluster corrected, 20 voxels, and $p < 0.05$. Figures corresponding to high gamma band were Bonferroni corrected to provide more stringent thresholding.

The beta band showed the greatest SIS, with large clusters in bilateral posterior temporal lobes. In left posterior temporal lobe, the very large cluster had two distinct peaks, one in posterior superior temporal gyrus, and one that extended through posterior inferior and middle temporal gyrus. The left hemisphere also showed a SIS peak in middle frontal gyrus, anterior to premotor cortex, cerebellum, and fusiform gyrus. The left regions of SIS present with unaltered

feedback were absent when speaking with altered auditory feedback. The right hemisphere had significant SIS in posterior temporal lobe with its peak in posterior middle temporal gyrus. This SIS cluster extended dorsally into right supramarginal gyrus. Right middle frontal gyrus, anterior to premotor cortex, showed SIS in a region analogous to that in the left hemisphere. SIS was in a second right frontal cluster, in superior frontal gyrus. The right hemisphere SIS clusters were absent when speaking with altered auditory feedback with the exception of a small cluster in the right cerebellum.

In the gamma band with unaltered feedback, left hemisphere SIS was in left precentral gyrus, left cerebellum, and in a large cluster around left posterior superior temporal gyrus that extends posteriorly through the posterior temporal lobe into a peak in angular gyrus, and further into a peak in precuneus. In the right hemisphere, there was a peak in posterior middle temporal gyrus, cerebellum, and a cluster in superior frontal gyrus very similar to the beta band cluster. In the left hemisphere, the SIS in the posterior temporal lobe and dorsally through the parietal lobe was absent with altered auditory feedback. Unexpectedly, the SIS in left frontal regions and in a cluster spanning left cerebellum into inferior temporal gyrus was much larger with 100 cent altered feedback. In the right hemisphere, the clusters in middle temporal gyrus and cerebellum did not have SIS with altered auditory feedback, while the cluster in right superior frontal gyrus showed SIS with 100 cent alteration, but not with 300 cent alteration. The high gamma band did not show SIS, but instead showed enhanced activity during speaking through frontal, motor and some parietal regions.

Peak Voxel	Region	T-value	p-value	SIS?	SIS w/ alt?
Theta/Alpha 4-13Hz					
Left					
-50.0 40.0 -10.0	inferior frontal gyrus	4.3228	2.4*10^-4	Anti-SIS	
-40.0 20.0 -25.0	anterior STG	3.2636	0.0051	Anti-SIS	
-45.0 -15.0 60.0	precentral gyrus	3.5841	0.0034	Anti-SIS	
-30.0 -45.0 70.0	postcentral gyrus	3.3236	0.0061	Anti-SIS	
-35.0 -50.0 25.0	temporal lobe/ sub-gyral	2.0835	0.0090	Anti-SIS	
Right					
45.0 -55.0 -25.0	cerebellum	-3.4812	2.4*10^-4	SIS	No
50.0 -55.0 -10.0	temporal lobe	-3.0809	0.0015	SIS	No
15.0 -45.0 70.0	postcentral gyrus	4.3416	4.9*10^-4	Anti-SIS	
45.0 -20.0 65.0	precentral gyrus	3.6632	0.0020	Anti-SIS	
10.0 30.0 60.0	superior frontal gyrus	2.7845	0.0044	Anti-SIS	
Beta 13-30Hz					
Left					
-50.0 5.0 50.0	middle frontal gyrus	-3.5565	0.0032	SIS	No
-60.0 -65.0 20.0	superior temporal gyrus	-3.0489	0.0020	SIS	No
-60.0 -60.0 -10.0	inferior/middle temporal gyrus	-3.1147	0.0095	SIS	No
-55.0 -35.0 -25.0	fusiform gyrus	-3.1096	9.8*10^-4	SIS	No
-55.0 -45.0 -30.0	cerebellum	-3.7240	2.4*10^-4	SIS	No
-60.0 -5.0 15.0	precentral gyrus	4.4002	0.0015	Anti-SIS	
-60.0 -15.0 30.0	postcentral gyrus	5.0901	4.9*10^-4	Anti-SIS	
-40.0 -30.0 65.0	postcentral gyrus	3.2286	0.0022	Anti-SIS	
-15.0 -50.0 70.0	postcentral gyrus	3.6527	7.3*10^-4	Anti-SIS	
Right					
60.0 -60.0 -35.0	cerebellum	-4.3149	2.4*10^-4	SIS	No
55.0 -70.0 15.0	posterior middle temporal gyrus/ occipital lobe	-2.1394	0.0254	SIS	No
60.0 -65.0 30.0	supramarginal gyrus	-2.5324	0.0137	SIS	No
40.0 10.0 55.0	middle frontal gyrus	-3.2860	0.0039	SIS	No
25.0 45.0 40.0	superior frontal gyrus	-3.1059	0.0037	SIS	No
15.0 -50.0 70.0	postcentral gyrus	3.1967	0.0063	Anti-SIS	
55.0 20.0 5.0	inferior frontal gyrus	2.7127	0.0073	Anti-SIS	
Gamma 30-55Hz					
Left					
-40.0 -5.0 30.0	precentral gyrus	-2.2136	0.0168	SIS	No
-65.0 -40.0 10.0	superior temporal gyrus	-2.5232	0.0095	SIS	No
-50.0 -75.0 30.0	temporal lobe/ angular gyrus	-2.2199	0.0117	SIS	No
-35.0 -75.0 45.0	parietal lobe/ precuneus	-2.8382	0.0022	SIS	No
-45.0 -15.0 60.0	precentral gyrus	3.1522	0.0142	Anti-SIS	
-40.0 -30.0 65.0	postcentral gyrus	3.2800	0.0042	Anti-SIS	
-45.0 -5.0 -40.0	inferior temporal gyrus	5.7502	2.4*10^-4	Anti-SIS	
-40.0 -60.0 -35.0	cerebellum	4.0000	4.9*10^-4	Anti-SIS	
Right					
35.0 -60.0 0.0	middle temporal gyrus	-3.4035	0.0020	SIS	No
50.0 -45.0 -45.0	cerebellum	-2.7119	0.0144	SIS	No
20.0 50.0 25.0	superior frontal gyrus	-2.4329	0.0173	SIS	Yes (100) No (300)
50.0 -5.0 -15.0	middle temporal gyrus	3.9903	0.0017	Anti-SIS	
65.0 5.0 5.0	superior temporal gyrus	3.1209	0.0068	Anti-SIS	
55.0 40.0 0.0	inferior frontal gyrus	2.3797	0.0107	Anti-SIS	

25.0 -30.0 65.0	postcentral gyrus	2.2375	0.0225	Anti-SIS	
65.0 -40.0 -10.0	middle temporal gyrus	3.0846	0.0024	Anti-SIS	
High Gamma 65-150Hz					
Left					
-55.0 -10.0 -30.0	fusiform gyrus	4.7832	2.4*10^-4	Anti-SIS	
-45.0 -20.0 30.0	postcentral gyrus	5.3715	2.4*10^-4	Anti-SIS	
-55.0 10.0 25.0	inferior frontal gyrus	6.6955	2.4*10^-4	Anti-SIS	
-5.0 45.0 -10.0	medial frontal gyrus	5.6544	4.9*10^-4	Anti-SIS	
-10.0 -40.0 70.0	postcentral gyrus	4.8536	4.9*10^-4	Anti-SIS	
Right					
65.0 -15.0 -10.0	middle temporal gyrus	4.9292	2.4*10^-4	Anti-SIS	
25.0 20.0 -25.0	anterior STG	4.2756	0.0015	Anti-SIS	
5.0 30.0 -35.0	rectal gyrus	5.2982	7.3*10^-4	Anti-SIS	
55.0 5.0 30.0	inferior frontal gyrus	4.7667	0.0020	Anti-SIS	

Table 3.1. Table summarizes information for the comparison between speaking and listening during unaltered auditory feedback. Table contains regions, peak voxel location, t-value, p-value, direction of modulation (SIS), and for regions showing SIS, presence of SIS during speaking with altered auditory feedback. For frequency bands theta/alpha, beta, and gamma results were cluster corrected, 20 voxels, $p<0.05$. For the high gamma band, results were Bonferroni corrected to provide more stringent thresholding.

3.3.4 Patterns of cortical activity during speaking change with altered auditory feedback.

In addition to the reduction in SIS through increased activity while speaking with altered auditory feedback, an extensive neural network showed enhanced responses while speaking with auditory feedback. All significant changes in neural activity while speaking with altered auditory feedback compared to speaking with unaltered feedback are summarized in Figure 3.4, Table 3.2. Regions that showed an increase in the absolute value of the power change when speaking with altered auditory feedback compared to unaltered auditory feedback are shown in the orange spectrum, while regions that showed decreased activity in auditory feedback while speaking with altered feedback compared to speaking with unaltered feedback are show with the cyan spectrum.

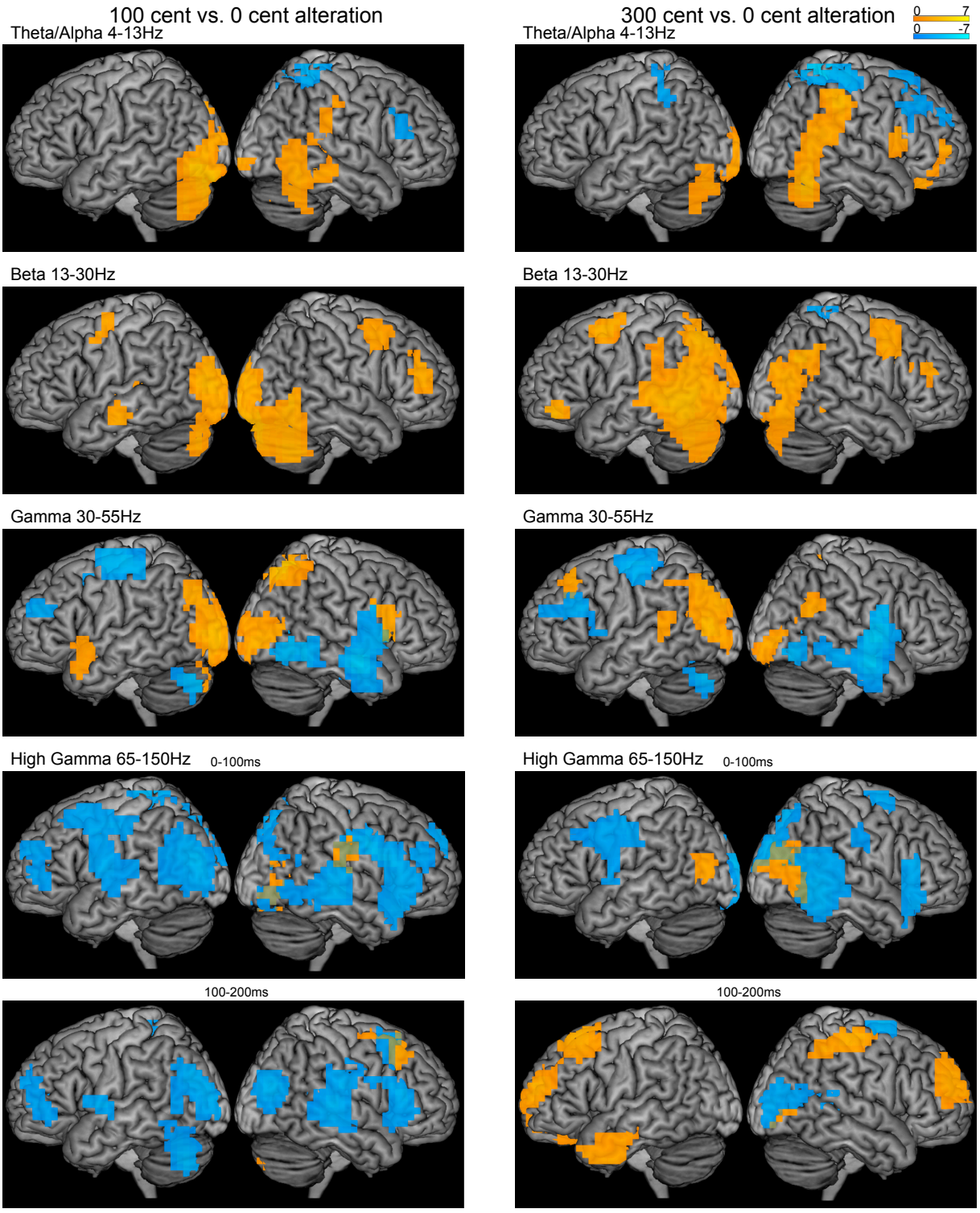

Figure 3.4. Cortical induced activity during speaking significantly changes when speaking during altered auditory feedback. Left column shows significant differences between speaking with 100 cent altered auditory feedback and unaltered auditory feedback. Right column shows significant differences between speaking with 300 cent altered auditory feedback and unaltered auditory feedback. Warm colors (orange spectrum) show regions with increased activity during altered auditory feedback compared to unaltered auditory feedback (positive alteration response) and cool colors (cyan spectrum) show regions with decreased activity during altered auditory feedback compared to unaltered auditory feedback (negative alteration response). All figures were cluster corrected, 20 voxels, and $p < 0.05$

In the theta/alpha band, positive alteration responses (increased activity with altered feedback) were seen in left occipital lobe and cerebellum to both 100 and 300 cent altered feedback. In the right hemisphere, regions of large regions of the posterior temporal lobe and dorsally into parietal cortex showed increased activity during speaking with both 100 cent and 300 cent altered feedback. In the 100 cent altered feedback the extent of the positive alteration response was smaller, with peaks in inferior and middle temporal gyrus and in inferior parietal lobe, while in the 300 cent altered feedback condition, the positive alteration responses in the right posterior temporal lobe were larger, with additional peaks in superior temporal gyrus extending to a large cluster through inferior parietal lobe into somatosensory cortex. The 300 cent altered feedback condition additionally showed two peaks in right inferior frontal gyrus with positive alteration responses.

Similarly, in the beta band several regions of cortex showed increased activity to altered auditory feedback, with larger responses to the 300 cent altered feedback. The left hemisphere shows positive alteration responses to the 100 cent altered feedback in precentral gyrus, occipital lobe, cerebellum, and in inferior/middle temporal gyrus. The positive alteration responses to the 300 cent altered feedback include an analogous region in precentral gyrus that extended to peak in middle frontal gyrus. In the 300 cent altered feedback, the left hemisphere positive alteration responses additionally included a large cluster in the posterior temporal lobe across superior and middle temporal gyrus that extended dorsally into the inferior parietal lobe and further into superior parietal cortex. In the 300 cent altered feedback condition an additional peak was also present in the ventral-anterior region of middle frontal gyrus. In right hemisphere, positive alteration responses to the 100 cent condition were seen in posterior temporal lobe extending into occipital lobe and cerebellum, and two clusters in middle frontal gyrus, one adjacent to precentral gyrus and one adjacent to inferior frontal gyrus-pars triangularis. The positive

alteration responses to the 300 cent condition included the two middle frontal gyrus regions and a large posterior temporal lobe cluster that extended dorsally into supramarginal gyrus.

In the gamma band, the patterns of positive alteration responses showed more variability between the 100 cent and 300 cent altered feedback. With the 100 cent altered feedback, positive alteration responses were in left inferior frontal gyrus and left precuneus, left cerebellum and left occipital lobe. In the right hemisphere, positive alteration responses were in the superior parietal lobe, occipital and inferior frontal gyrus. In contrast, in the 300 cent altered feedback condition, regions in left hemisphere showing positive alteration responses were superior temporal gyrus, posterior middle temporal gyrus, inferior parietal lobe, superior parietal lobe, occipital lobe and middle frontal gyrus. In the right hemisphere clusters that extends from superior temporal gyrus to supramarginal gyrus, middle occipital gyrus and precuneus showed positive alteration responses.

In the high gamma band, a large extent of speech motor cortex showed decreased activity when speaking with altered feedback compared to speaking with unaltered feedback, but a few regions show increased activity to altered auditory feedback. In the 100 cent altered feedback condition, positive alteration responses were only in right postcentral gyrus, right posterior middle temporal gyrus and right middle frontal gyrus. In the 300 cent altered feedback condition, several more regions show positive alteration responses. In the left hemisphere, positive alteration responses were in posterior middle temporal gyrus, anterior temporal lobe, two clusters in superior frontal gyrus and a small cluster in rectal gyrus. In the right hemisphere, positive alteration responses were in posterior temporal lobe, inferior parietal lobe, postcentral gyrus, middle frontal gyrus and superior frontal gyrus.

Peak Voxel	Region	T-value	p-value	Alteration response?
Theta/Alpha 4-13Hz				
100cent vs. 0cent				
Left				
-45.0 -80.0 -10.0	inferior occipital gyrus	4.2842	7.3*10^-4	positive
-55.0 -70.0 -45.0	cerebellum	3.4151	0.0017	positive
Right				
20.0 -100.0 -5.0	cuneus	2.5689	0.0137	positive
45.0 -55.0 -15.0	fusiform gyrus	2.8853	0.0022	positive
55.0 -35.0 -15.0	middle temporal gyrus	2.4911	0.0112	positive
55.0 -30.0 25.0	inferior parietal lobe	2.1954	0.0154	positive
55.0 35.0 20.0	middle frontal gyrus	-2.3536	0.0164	negative
25.0 -45.0 65.0	superior parietal lobe	-3.1653	0.0100	negative
300cent vs. 0cent				
Left				
-55.0 -75.0 -40.0	cerebellum	2.1503	0.0237	positive
-15.0 -100.0 -10.0	occipital lobe/lingual gyrus	2.5176	0.0049	positive
-50.0 -45.0 50.0	inferior parietal lobe	-2.5492	0.0261	negative
Right				
45.0 -60.0 -25.0	cerebellum	3.7525	4.9*10^-4	positive
50.0 -55.0 5.0	middle temporal gyrus	2.8081	4.9*10^-4	positive
50.0 -45.0 15.0	superior temporal gyrus	2.8424	0.0037	positive
60.0 -30.0 45.0	postcentral gyrus	3.5131	7.3*10^-4	positive
45.0 55.0 5.0	inferior frontal gyrus	2.6337	0.0115	positive
45.0 20.0 15.0	inferior frontal gyrus	2.6077	0.0046	positive
45.0 20.0 35.0	middle frontal gyrus	-2.5115	0.0161	negative
15.0 15.0 45.0	cingulate gyrus	-2.9089	0.0063	negative
20.0 50.0 30.0	superior frontal gyrus	-2.4599	0.0159	negative
15.0 -50.0 70.0	postcentral gyrus	-4.7517	4.9*10^-4	negative
Beta 13-30Hz				
100cent vs. 0cent				
Left				
-60.0 -15.0 -20.0	inferior/middle temporal gyrus	2.6717	0.0093	positive
-45.0 0.0 50.0	precentral gyrus	2.6017	0.0066	positive
-35.0 -85.0 10.0	middle occipital gyrus	3.5259	0.0037	positive
-30.0 -80.0 -35.0	cerebellum	3.0162	0.0034	positive
Right				
15.0 -75.0 -10.0	lingual gyrus	4.1615	7.3*10^-4	positive
50.0 -60.0 -45.0	cerebellum	3.2483	0.0044	positive
60.0 -60.0 -5.0	middle temporal gyrus	3.3480	0.0020	positive
30.0 10.0 40.0	sub-gyral/middle frontal gyrus/precentral	2.6264	0.0127	positive
45.0 45.0 15.0	middle frontal gyrus	2.8661	0.0081	positive
300cent vs. 0cent				
Left				
-45.0 45.0 -15.0	middle frontal gyrus	3.4635	7.3*10^-4	positive
-45.0 5.0 55.0	middle frontal gyrus	3.1937	0.0017	positive
-60.0 -35.0 35.0	inferior parietal lobe	2.1957	0.0200	positive
-55.0 -50.0 -5.0	temporal lobe	3.8117	0.0017	positive
-45.0 -60.0 20.0	superior temporal gyrus	2.9483	0.0088	positive
-30.0 -70.0 45.0	superior parietal lobe	3.3717	0.0024	positive
-60.0 -65.0 0.0	posterior middle temporal gyrus	3.7148	0.0061	positive
-50.0 -85.0 5.0	middle occipital gyrus/ba19	2.8270	0.0081	positive
-35.0 -60.0 -30.0	cerebellum	2.8583	0.0149	positive

Right				
60.0 -65.0 30.0	supramarginal gyrus	3.6809	0.0029	positive
40.0 -75.0 15.0	middle occipital gyrus	2.8376	0.0068	positive
40.0 -80.0 -10.0	inferior occipital gyrus	2.6796	0.0156	positive
40.0 -85.0 -40.0	cerebellum	2.0950	0.0278	positive
40.0 -50.0 -5.0	temporal lobe	2.4710	0.0085	positive
35.0 5.0 40.0	precentral gyrus	2.9238	0.0024	positive
55.0 40.0 25.0	middle frontal gyrus	2.6336	0.0105	positive
15.0 -50.0 70.0	postcentral gyrus	-2.5492	0.0129	negative
Gamma 30-55Hz				
100cent vs. 0cent				
Left				
-30.0 10.0 -15.0	inferior frontal gyrus	2.3589	0.0139	positive
-35.0 -75.0 35.0	precuneus	2.8716	0.0117	positive
-35.0 -95.0 5.0	middle occipital gyrus	3.8549	2.4*10^-4	positive
-15.0 -80.0 -40.0	cerebellum	2.6593	0.0115	positive
-55.0 -80.0 -30.0	cerebellum	-4.0998	7.3*10^-4	negative
-25.0 50.0 25.0	middle frontal gyrus	-2.8296	0.0015	negative
-40.0 -20.0 65.0	precentral gyrus	-3.5152	0.0049	negative
Right				
45.0 -90.0 -5.0	inferior occipital gyrus	3.1340	0.0061	positive
20.0 -65.0 55.0	superior parietal lobe	5.1235	2.4*10^-4	positive
60.0 15.0 0.0	inferior frontal gyrus	3.0007	0.0027	positive
65.0 -50.0 -10.0	middle temporal gyrus	-3.5784	0.0037	negative
45.0 -60.0 -10.0	occipital lobe	-3.5210	0.0039	negative
45.0 -5.0 -20.0	middle temporal gyrus	-4.5224	2.4*10^-4	negative
65.0 5.0 25.0	inferior frontal gyrus	-2.7995	0.0210	negative
300cent vs. 0cent				
Left				
-65.0 -40.0 10.0	superior temporal gyrus	2.3031	0.0298	positive
-45.0 -45.0 45.0	inferior parietal lobe	1.9094	0.0376	positive
-35.0 -75.0 45.0	superior parietal lobe	3.0417	0.0071	positive
-50.0 -80.0 10.0	posterior middle temporal gyrus/occipital lobe	3.5435	9.8*10^-4	positive
-30.0 30.0 50.0	middle frontal gyrus	2.5653	0.0125	positive
-55.0 -80.0 -30.0	cerebellum	-3.8054	9.8*10^-4	negative
-35.0 -20.0 55.0	precentral gyrus	-3.6613	4.9*10^-4	negative
-30.0 45.0 20.0	superior frontal gyrus	-2.1216	0.0093	negative
-55.0 25.0 25.0	inferior frontal gyrus	-2.8807	0.0054	negative
-60.0 5.0 5.0	anterior superior temporal gyrus	-2.0555	0.0464	negative
Right				
45.0 -90.0 0.0	middle occipital gyrus	2.7196	0.0090	positive
0.0 -40.0 45.0	precuneus	2.4891	0.0098	positive
50.0 -55.0 20.0	superior temporal gyrus/ supramarginal gyrus	2.5798	0.0144	positive
60.0 -5.0 -20.0	inferior/middle temporal gyrus	-4.2306	4.9*10^-4	negative
65.0 -5.0 0.0	superior temporal gyrus	-4.4599	0.0020	negative
65.0 5.0 20.0	precentral gyrus	-3.5064	0.0068	negative
40.0 -70.0 -10.0	inferior occipital gyrus	-3.2882	9.8*10^-4	negative
65.0 -30.0 -15.0	middle temporal gyrus	-2.7581	0.0251	negative
High Gamma 65-150Hz				
100cent vs. 0cent				
Left				
0-100ms				
-30.0 55.0 25.0	middle frontal gyrus	-2.4734	0.0127	negative

-50.0 55.0 5.0	inferior frontal gyrus	-2.5797	0.0015	negative
-60.0 0.0 40.0	precentral gyrus	-2.7136	0.0225	negative
-55.0 -10.0 45.0	precentral gyrus	-2.6503	0.0247	negative
-60.0 -55.0 0.0	middle temporal gyrus	-2.3444	0.0229	negative
-50.0 -80.0 20.0	posterior middle temporal gyrus	-3.4150	0.0017	negative
100-200ms				
-65.0 -5.0 5.0	superior temporal gyrus	-2.8415	0.0186	negative
-50.0 -70.0 -40.0	cerebellum	-3.3144	0.0029	negative
Right				
0-100ms				
65.0 -20.0 20.0	postcentral gyrus	1.9590	0.0090	positive
35.0 -70.0 15.0	posterior middle temporal gyrus/ middle occipital gyrus	1.7499	0.0359	positive
10.0 -75.0 45.0	precuneus	-2.7642	0.0112	negative
45.0 -75.0 -10.0	inferior occipital gyrus	-1.7828	0.0256	negative
65.0 -25.0 -5.0	middle temporal gyrus	-3.1431	0.0044	negative
60.0 5.0 25.0	inferior frontal gyrus	-3.5647	0.0076	negative
60.0 15.0 20.0	inferior frontal gyrus	-3.2990	0.0049	negative
45.0 20.0 -20.0	superior temporal gyrus	-2.6761	0.0247	negative
50.0 45.0 15.0	inferior frontal gyrus	-1.6979	0.0217	negative
5.0 55.0 40.0	medial frontal gyrus	-2.6504	0.0090	negative
50.0 25.0 10.0	inferior frontal gyrus	-2.7822	0.0205	negative
100-200ms				
35.0 25.0 50.0	middle frontal gyrus	2.5550	0.0100	positive
65.0 -25.0 5.0	superior temporal gyrus	-3.1093	0.0110	negative
30.0 -75.0 20.0	temporal lobe	-3.2400	0.0037	negative
70.0 -20.0 20.0	postcentral gyrus	-2.9997	0.0022	negative
300cent vs. 0cent				
Left				
0-100ms				
-55.0 -75.0 5.0	posterior middle temporal gyrus/ occipital lobe	2.3870	0.0283	positive
-55.0 5.0 40.0	precentral gyrus	-2.3693	0.0203	negative
-55.0 -20.0 40.0	postcentral gyrus	-2.2589	0.0242	negative
-60.0 0.0 0.0	superior temporal gyrus	-2.3020	0.0339	negative
-5.0 -100.0 15.0	cuneus	-3.6563	0.0012	negative
100-200ms				
-30.0 60.0 20.0	superior frontal gyrus	3.1029	0.0029	positive
-35.0 25.0 55.0	superior frontal gyrus	2.8500	0.0066	positive
-10.0 20.0 -30.0	rectal gyrus	2.7956	0.0159	positive
-50.0 0.0 -40.0	inferior temporal gyrus	2.6726	0.0100	positive
-40.0 -15.0 -15.0	temporal lobe	2.0954	0.0437	positive
Right				
0-100ms				
40.0 -65.0 10.0	posterior temporal lobe	2.2305	0.0039	positive
25.0 -85.0 30.0	cuneus	-2.9740	0.0039	negative
45.0 -50.0 15.0	superior temporal gyrus	-2.8777	2.4*10^-4	negative
60.0 -50.0 -5.0	middle temporal gyrus	-3.1245	9.8*10^-4	negative
60.0 -20.0 35.0	postcentral gyrus	-2.0159	0.0149	negative
60.0 30.0 15.0	inferior frontal gyrus	-2.7718	0.0242	negative
50.0 25.0 0.0	inferior frontal gyrus	-2.5500	0.0181	negative
30.0 20.0 -30.0	anterior superior temporal gyrus	-2.0174	0.0173	negative
100-200ms				
50.0 -45.0 60.0	inferior parietal lobe	2.7747	0.0107	positive

45.0 -20.0 60.0	postcentral gyrus	2.9899	0.0059	positive
50.0 50.0 15.0	middle frontal gyrus	1.7649	0.0225	positive
20.0 55.0 20.0	superior frontal gyrus	2.8485	0.0068	positive
5.0 0.0 65.0	medial frontal gyrus	-3.3266	0.0083	negative
40.0 -85.0 5.0	middle occipital gyrus	-2.6219	0.0068	negative

Table 3.2. Table summarizes information for the comparison between speaking during altered auditory feedback and speaking during unaltered auditory feedback. Table contains regions, peak voxel location, t-value, p-value, and direction of alteration response. All results were cluster corrected, 20 voxels, and $p < 0.05$.

3.3.5 Regions that show Speaking Induced Suppression are largely overlapping with regions that show Positive Alteration Responses.

The majority of cortical regions that showed SIS also showed significantly increased responses to speaking with altered auditory feedback compared with speaking with unaltered auditory feedback. The overlay of regions that showed both SIS and positive alteration responses are shown in figure 3.5, left panel. The overlapping regions of SIS and positive alteration responses are shown in green. In both the theta/alpha band and the beta band nearly the full extent of clusters with SIS also showed positive alteration responses. In the gamma band, while large regions that showed SIS also showed positive alteration responses, a few regions, including left precentral gyrus, left cerebellum, right posterior temporal lobe and superior frontal gyrus showed SIS but do not show positive alteration responses.

Some regions that showed anti-SIS also showed a significant decrease in responses to speaking with altered auditory feedback compared with speaking with unaltered auditory feedback. The overlay of regions that show both ant-SIS and negative alteration responses are shown in figure 3.5, right panel with the overlapping regions in purple. In the theta/alpha band small regions of bilateral postcentral gyrus showed overlapping anti-SIS and negative alteration responses. In the right hemisphere the cluster extends into superior parietal lobe. In the beta band only a small cluster in right dorsal postcentral gyrus into superior parietal lobe showed overlapping anti-SIS and negative alteration responses. In the gamma band, anti-SIS and

negative alteration responses overlapped in left precentral and postcentral gyri and left cerebellum. In the right hemisphere, anti-SIS and negative alteration responses overlapped throughout the temporal lobe and into inferior frontal gyrus. In the high gamma band there was a large overlap between anti-SIS and negative alteration responses, including clusters throughout a large region of left precentral gyrus, middle and inferior frontal gyri and in right middle temporal gyrus, anterior superior temporal gyrus and inferior frontal gyrus.

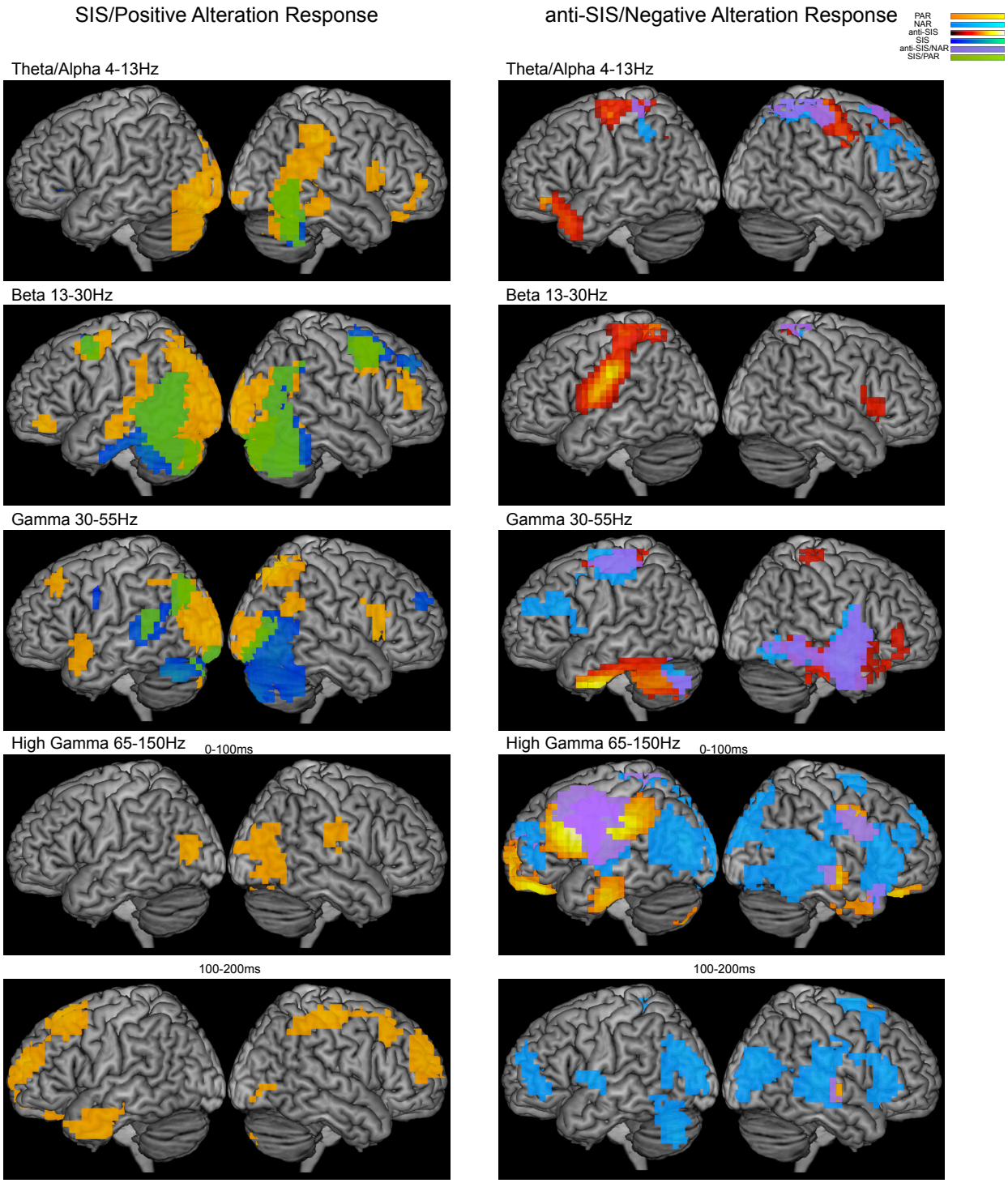

Figure 3.5. Overlay between significant changes between speaking and listening with unaltered auditory feedback and between speaking during altered auditory feedback compared to unaltered auditory feedback. Left column shows overlay of regions that show SIS (blue) and positive alteration responses (orange). Overlapping regions are shown in green. Right column shows overlay of regions that show anti-SIS (red-yellow) and negative alteration responses (cyan). Overlapping regions are shown in purple.

3.4 Discussion

In this study we examined the distribution of SIS across frequency bands and the cortical surface, and how cortical activity is modulated during speaking when auditory feedback deviates by either a small or large degree from the expected feedback. We showed that SIS is present in the alpha/theta, beta and gamma bands and, with a few exceptions, is eliminated with pitch altered auditory feedback. We showed that this absence of SIS is driven by a significant increase in neural response when speaking with altered feedback. Cortical networks in both hemispheres show increased activity to pitch-altered auditory feedback in all frequency bands. The networks showing positive alteration responses were sensitive to the degree of alteration, showing greater responses when the auditory feedback had a considerable deviation from expectation (300 cent alteration) than when the auditory feedback had a small deviation from expectation (100 cent alteration).

In this study, we showed that SIS extends beyond the temporal lobe with regions in the frontal lobe, parietal lobe and cerebellum showing suppression behavior in addition to large regions of the posterior temporal lobe. The largest extent of SIS throughout the cortex occurred in auditory and frontal regions in the beta band. The only previous work to explore the effect of frequency band on SIS compared low-frequency evoked responses to high gamma induced responses using ECoG (Greenlee et al., 2011), and found that evoked SIS shows a distinct spatial pattern from high gamma induced SIS. In the current study, we confirmed the considerable spatial overlap in SIS across frequency bands, however in addition we show that several SIS regions, mostly frontal and parietal, vary across frequency bands. In our study we do not see SIS in the high gamma band, and notably, we see neither suppression nor enhancement in bilateral temporal lobes. This is in direct contrast with the Greenlee et al. study, which found both SIS and anti-SIS, often separated only be millimeters, in the posterior temporal lobe. The absence of

speaking-induced modulation in the current study is likely due to a combination of 1) the inferior spatial resolution of MEG compared to ECoG and 2) both the smoothing kernel and the stringent statistical thresholding of a 20 voxel cluster causing the superposition of SIS and anti-SIS.

In this study, we explored how SIS is modulated in the presence of altered auditory feedback, causing a mismatch between the expectation of the auditory feedback and the perceived auditory consequence of speech. In replication of previous results (Behroozmand and Larson, 2011; Heinks-Maldonado et al., 2006), we demonstrated that evoked SIS calculated from temporal electrodes is reduced and eliminated when the auditory feedback is altered. We further showed that cortical regions had reduced and eliminated SIS with altered auditory feedback, primarily driven by a significant increase in neural responses to speaking with altered auditory feedback. These results are in accord with the single unit recordings in non-human primates. That work found neurons showing suppression at vocalization onset, showed reduced suppression at the onset of vocalization when the auditory feedback is shifted by 200 cents (Eliades and Wang, 2008). This work goes on to interpret this finding as evidence that the altered feedback response combines with the vocalization-induced suppression. Interestingly, while that study reports an overall bias of neurons that showed suppression to show less suppression with altered auditory feedback, they also report that some neurons that did not show suppression to natural vocalizations show decreased activity to the altered auditory feedback. This pattern was also true in our current study across all frequency bands, with regions showing anti-SIS correspondingly showing negative alteration response. Cortical responses to altered auditory feedback showed enhancement during speaking beyond those clusters with SIS. This network of enhanced responses to altered feedback is larger as the perceived auditory feedback deviates further from the expected auditory consequences of vocalization. Taken together, there

exists a true reduction of SIS when the perceived auditory feedback does not match the expected auditory feedback. Additionally, a larger network shows increased activity that is modulated with the amount of deviation in the perceived auditory feedback from the expected auditory feedback.

In this study, subjects did not compensate to the altered auditory feedback which means that the changes in cortical responses during speaking under altered feedback must be driven by the perception of the altered feedback and the recognized mismatch between the expected and perceived feedback, rather than the initiation of a motor change. The absence of compensation is likely due to experimental design- the combination of the randomly interleaved levels and direction of altered feedback and of the subjects producing short, ballistic utterances. In this environment, subjects are continuing to monitor their auditory feedback, but are not using the feedback to alter their motor plan. In this way, the cortical responses reported in this study reflect feedback monitoring.

3.5 Conclusions

This study aimed to describe speech-induced suppression across the cortex and frequency bands. We confirmed previous findings of SIS in the temporal lobes and extended these findings to demonstrate similar SIS response profiles in frontal and cerebellar regions. We found both the largest extent and the greatest magnitude of suppression during speaking in the beta band. Regions that showed SIS with unaltered auditory feedback showed reduced or eliminated SIS when speaking with altered auditory feedback. The elimination of SIS was driven by increased activity in the same regions during speaking with altered auditory feedback. Furthermore, a larger network showed increased activity that was modulated with the amount of deviation in the perceived auditory feedback from the expected auditory feedback.

Chapter 4.

Inter-hemispheric communication coordinates vocal feedback control of pitch during ongoing phonation.

4.1 Introduction

Vocal control of pitch is essential during vocal communication in humans, providing semantic cues through voice onset and conveying semantic and affective prosody. Not surprisingly, therefore, pitch production has been an active field of study. Despite the number of studies, the understanding of sensory-motor control of pitch production is confounded by contradictory findings resulting from the use of inadequate techniques and the inability to address behavioral variability across subjects. The majority of work was conducted with EEG (Behroozmand et al., 2011b; 2011a), which lacks the spatial resolution. Work has also been done in ECoG, which, being limited by grid coverage, primarily reported sensory-motor results in the left hemisphere (Chang et al., 2013; Greenlee et al., 2013) while work in fMRI lacks the temporal resolution to distinguish a sensory-motor event, and reported bilateral superior temporal gyrus (Parkinson et al., 2012). A single study using magnetoencephalography (MEG) by our lab did find bilateral sensory-motor responses during vocal control of pitch, but this study used a targeted ROI approach and the evoked response, unlike the high gamma ECoG findings, did not correlate with the subjects' motor behavior (Kort et al., 2014). In order to truly address the sensory-motor responses during vocal control of pitch, both space and time resolved neuroimaging must be conducted. The high time and spatial resolution of MEG imaging of the high gamma band- reflecting spiking activity of neural groups- provides this resolution.

Here we aim to address the following hypotheses: 1) both the left and the right hemisphere are involved in feedback control of speech and 2) inter-hemispheric communication

coordinates vocal feedback control. Measuring the neural and behavioral response to an unexpected shift in pitch allows the investigation of how the brain responds to unexpected reafferent information. Using this approach, we sought to investigate which cortical regions are involved in voice pitch control when removed from linguistic and emotional context, how these cortical responses to an error in pitch production evolve over time, how these cortical responses relate to behavioral responses, and the neural connectivity between nodes in the pitch production network. In order to address these questions we used MEG to obtain whole-brain cortical responses to a brief shift in pitch feedback. We found left sensory regions respond rapidly to detect an unexpected error in pitch, while right temporal, parietal, premotor and frontal regions show both a rapid and sustained response monitoring feedback and initiating the behavioral motor response.

4.2 Materials and Methods

4.2.1 Subjects

Twelve right-handed English speaking volunteers with normal speech and hearing (6 female) subjects showed behavioral compensation the auditory perturbation and were included in all analyses. After procedures had been fully explained all participants gave their informed consent. The study was performed with the approval of the University of California, San Francisco Committee for Human Research.

4.2.2 MEG Recording

The task was completed during whole head MEG neural recording in awake subjects laying in the supine position. The MEG system (MISL, Coquitlam, British Columbia, Canada) consists of 275 axial gradiometers and was recorded with a sampling rate of 1200Hz. Three

fiducial coils were placed on the nasion and left/right preauricular points to triangulate the position of the head relative to the MEG sensor array. In a separate session high resolution anatomical MRIs were obtained for each subject. The fiducial markers points were later co-registered with an anatomical MRI to generate head shape.

4.2.3 Experimental Design and Procedure

A full description of the experimental design and procedure is available in section 2.2.3. To summarize, the experiment was administered in a block design with a Speaking Condition and Listening Condition. In the Speaking Condition, subjects were instructed to speak the vowel /a/ until the termination cue. The subjects spoke into an MEG-compatible optical microphone and received auditory feedback through earplug earphones. During the phonation the subjects heard one 100-cent pitch perturbation lasting 400ms whose onset was jittered in time from speech onset. An equal number of pitch shifts that either raised or lowered the perceived pitch were pseudorandomly distributed across the experiment. In the Listening Condition, subjects received the same visual prompts but passively listened to the recording of their perturbed voice feedback obtained in the previous Speaking Condition block. The auditory input through earphones in both conditions was identical. The blocks alternated between the two conditions: blocks 1 and 3 were the Speaking Condition; blocks 2 and 4 were the Listening Condition.

4.2.4 Audio analysis

The pitch analysis was performed on each subject's single trial audio data. The details of the pitch analysis are in section 2.2.4.

4.2.5 MEG data analysis

The NUTMEG time-frequency spatially adaptive filter algorithm (Dalal et al., 2008) was used to localize the induced activity in the high gamma band (65-150Hz) to the individual subject's spatially normalized MRI's. The NUTMEG algorithm has been previously described in detail in Dalal et al, 2011 (Dalal et al., 2011). Noise corrected pseudo-F ratios were computed between the active windows (following the perturbation onset) and the prestimulus control baseline (the window preceding the onset of the perturbation). The windows were 100ms with an overlap of 25ms, allowing the reconstruction resolution of 100ms following the perturbation onset to 300ms following the perturbation onset. Group statistics were computing using the NUTMEG time-frequency statistics toolbox (Dalal et al., 2011; 2008) with statistical non-parametric mapping (SnPM) (Singh et al., 2003). A brief description of this method is available in section 3.2.6.2.

4.2.6 Functional Connectivity Analysis

The Phase Locking Value (PLV) approach (Lachaux et al., 1999; Mormann et al., 2000) was used to evaluate the dynamics of the phase synchrony among the nodes in the speech motor network. The speech-motor network was identified as the 13 regions showing a significant power increase during the speaking condition in response to the unexpected pitch shift (Table 4.2) along with left hemisphere analogs of the peaks in IFG- pars orbitalis, IFG- pars triangularis, MFG and PMC. Left hemisphere premotor and frontal nodes are often implicated in speech tasks, so they were included in this analysis. Time courses of source intensities were computed with CTF software tools using an adaptive spatial filtering technique with using a signal bandwidth of 0-300 Hz. Single trial source time-courses were then filtered 65-150Hz. Their instantaneous phases were extracted via a Hilbert Transform (dismissing 0.1 seconds at each end to avoid the border

effects). From this time-course data, PLV was calculated in overlapped time windows of 100 ms, in 10ms steps. If $\theta(t,n)$ is the phase difference for the time window centered at t in trial n, the PLV will be calculated as the average value:

$$PLV_t = \frac{1}{N}\left|\sum_{n=1}^{N}\exp(i\theta(t,n))\right|$$

PLV was calculated for a baseline period during sustained speech (not including speech onset) but preceding the unexpected pitch shift and during the time of the pitch shift. Percent change from baseline was calculated. To control the family-wise error due to multiple comparisons, every Wilcoxon signed rank test was subjected to a permutation test (Ernst, 2004). This was accomplished by randomly dividing the participants into two sets, matching the numbers in the original groups. The Wilcoxon test was then carried out in these two new groups. This procedure was repeated 2000 times and the p-value from each test was retained in order to obtain a p-value distribution. We then identified the 5th percentile of each distribution, and only p-values below that threshold are considered significant. We then applied a time-clustering correction, requiring significant PLV changes to be sustained for a minimum of two consecutive time windows.

The PLV time windows were collapsed to three time periods- prior to compensation onset including windows centered 50-130ms, spanning the time 0-180ms following the pitch shift; during compensation onset including windows centered 140-220ms, spanning the time 90-270ms; following compensation onset including windows centered 230-350ms, spanning the time 180-400ms, by calculating the mean of significant changes in PLV.

4.3 Results

4.3.1 Variable compensation to induced pitch shifts.

Subjects were instructed to produce sustained utterances of the vowel /a/ into a microphone and hear their auditory feedback through headphones during the speaking condition during MEG recording. In each utterance, a 400ms pitch shift was applied in real time with a digital signal processor, pseudorandomly alternating between increasing and decreasing the pitch of the subject's voice by 100 cents (1/12 octave). Subjects were not given instructions regarding the pitch shift or whether to respond. In a subsequent block, the listening condition, subjects listened to the playback of their voice recorded during the speaking condition, including the unexpected pitch shift. Subjects responded to the transient pitch shift by rapidly changing their pitch production to oppose shifts that either increased or decreased the pitch of their voice. Changing one's vocalization to oppose the direction of the feedback alteration is termed compensation. The mean f0 contours of each subject in response to the pitch shift are shown in figure 4.1. The mean compensation was 21.79% (range: 10.4-41.65). Subjects also showed variability across trials, with a mean standard deviation of 32.02 (range: 20.30- 57.63). The mean compensation onset was 187.22ms (range: 124.69-300.29), while the mean peak of compensation was 522.94ms (range: 458.48-625.37). Individual subject peak compensation, latency and variability are summarized in table 4.1. There were no correlations between percent compensation, variability, compensation onset latency, and compensation peak latency across subjects.

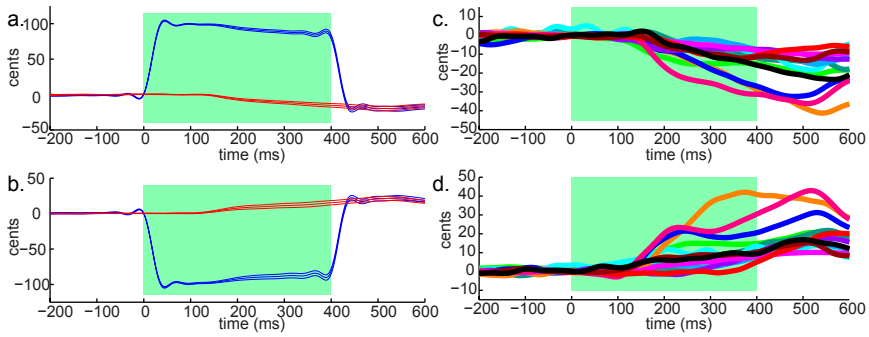

Figure 4.1. Vocal responses to the shift in pitch of audio feedback. Individual subject mean responses to the +100 cent, upper plots (a, c), and the -100 cent (b, d), lower plots pitch shift. The shift onset occurs at 0ms and is sustained for 400ms, denoted by the green region.
a, b. Grand-average of vocal responses across subjects. Blue traces show the mean time course of feedback heard by subjects, the thick blue line is the grand-average over subjects and flanking thin blue lines are +/- standard errors. In a similar fashion, red traces show the mean time course of the pitch produced by subjects. c,d. Individual subjects' mean time courses of the produced pitch.

Subject Number	Mean Percent Compensation	Variability of Compensation	Mean Compensation Onset (ms)	Mean Compensation Peak (ms)
1	19.91	25.13	211.77	573.13
2	20.80	30.35	124.69	519.43
3	16.04	57.63	146.46	523.79
4	15.32	34.86	256.76	525.24
5	41.65	31.59	178.39	458.48
6	31.83	34.36	191.45	506.37
7	15.70	39.66	300.29	625.37
8	10.04	20.30	156.62	525.24
9	39.54	28.13	130.50	513.63
10	16.29	28.89	253.85	464.29
11	14.23	26.92	140.66	507.82
12	20.11	26.44	155.17	532.49
Mean (std)	**21.79 (10.25)**	**32.02 (9.52)**	**187.22 (56.95)**	**522.94 (44.02)**

Table 4.1. Table contains behavioral data on subject's amount of compensation, variability of compensation, latency of compensation onset, and latency of peak compensation.

4.3.2 Left sensory cortex responses direct pitch shift detection.

While speaking with an externally applied pitch shift, within the first 125ms after pitch shift onset we observed widespread increases in induced high gamma power (HGP) bilateral over left primary and secondary sensory regions and right premotor, parietal and frontal regions, following which HGP enhancement becomes right dominant (Fig 4.2, Table 4.2). In the left hemisphere, primary sites were in sensory regions across auditory cortices in the temporal lobe, and dorsally including somatosensory cortex (SSC). In the right hemisphere, activations included

the inferior parietal lobe (IPL) including the supramarginal gyrus (SMG), right premotor cortex (PMC), and anterior regions including the insula and inferior frontal gyrus (IFG). The timing of the activity (i.e. HGP increases) indicates the involvement of these regions in the first stages of pitch feedback error detection.

Cortical regions that show activity that peaks after the onset of the pitch shift, yet prior to the onset of behavioral compensation (187.22ms), are presumably involved in detecting the error in feedback, processing that error and preparing for a behavioral change. The left hemisphere primary and secondary sensory regions, with the exception of left posterior middle temporal gyrus (MTG), had peak activity during this detection, processing and preparation window. In contrast, only two regions in the right hemisphere showed this profile of early activity peaks, right IFG- pars triangularis and anterior superior temporal gyrus (STG). The peak in right anterior STG was significant and reported in table 4.2, but did not survive the more stringent thresholding in figure 4.2.

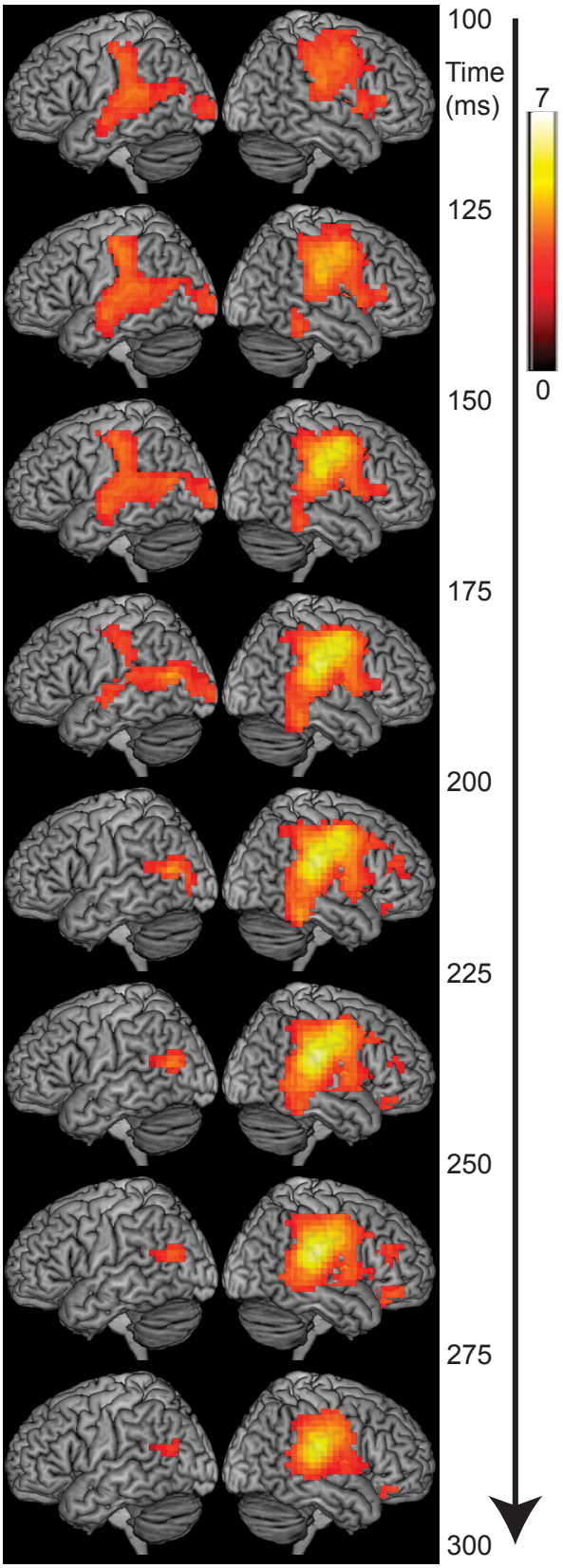

Figure 4.2. MEG cortical responses are enhanced in the speaking condition in response to the pitch shift compared to steady state vocalization. Images are 20 voxels cluster corrected, $p<0.005$.

Region	Perturbed Speech-Speech MNI peak voxel	Perturbed Speech-Speech Time-course	T-value	P-value
Left Transverse Temporal Gyrus	-45.0 -30.0 10.0	Peak: 100-125ms Duration: 100-200ms	3.96685	2.4*10^-4
Left Somatosensory Cortex	-55.0 -25.0 40.0	Peak: 125-150ms Duration: 100-200ms	4.02157	9.8*10^-4
Left Posterior Middle Temporal Gyrus	-60.0 -70.0 15.0	Peak: 175-200ms Duration: 100-300ms	4.54941	4.9*10^-4
Left Middle Occipital Lobe	-10.0 -100.0 10.0	Peak: 125-150ms Duration: 100-150ms	3.34456	0.0081
Left Middle Temporal Gyrus	-65.0 -10.0 -5.0	Peak: 175-200ms Duration: 100-250ms	3.83291	0.0012
Right Premotor Cortex	55.0 -10.0 40.0	Peak: 175-200ms Duration: 100-300ms	5.57166	2.4*10^-4
Right Inferior Parietal Lobe	55.0 -30.0 40.0	Peak: 175-200ms Duration: 100-300ms	4.49757	4.9*10^-4
Right Supramarginal Gyrus	55.0 -25.0 20.0	Peak: 225-250ms Duration: 100-300ms	6.35952	2.4*10^-4
Right Middle Temporal Gyrus	60.0 -40.0 -15.0	Peak: 200-225ms Duration: 100-275ms	3.63265	9.8*10^-4
Right Insula	40.0 0.0 5.0	Peak: 250-275ms Duration: 100-275ms	4.56657	4.9*10^-4
Right Inferior Frontal Gyrus-pars triangularis	60.0 30.0 10.0	Peak: 100-125ms Duration: 100-200ms	3.49622	0.002
Right Anterior Superior Temporal Gyrus	65.0 10.0 -5.0	Peak: 150-175ms Duration: 100-275ms	3.62678	0.0017
Right Middle Frontal Gyrus	60.0 30.0 25.0	Peak: 200-225ms Duration: 200-275ms	3.27512	0.0068
Right Inferior Frontal Gyrus-pars orbitalis	30.0 35.0 -10.0	Peak: 250-275ms Duration: 200-300ms	3.52064	0.0015

Table 4.2. Table contains regions, peak voxel location, duration, peak latency, t-value and p-value of regions of significant enhancement in response to the pitch shift during speaking compared to steady state vocalization. 20 voxels cluster corrected, $p<0.01$.

4.3.3 Right premotor and parietal cortex, left posterior temporal cortex initiate compensation and prepare for change in feedback.

Cortical regions that show peak activity concurrent with the compensation onset suggest involvement in both inducing compensation and preparing for the new feedback that will result from the motor change. The onset of compensation occurs in the window 175-200ms following the pitch shift. During this window the HGP continued to increase in the right hemisphere with peaks in right premotor cortex and right dorsal IPL, while the HGP response in the left hemisphere was restricted to left posterior MTG.

4.3.4 Right frontal and parietal regions monitor the feedback change feedback.

As subjects change their pitch to compensate for the shift, their auditory feedback also changes. Cortical regions that show peak activity after the onset of compensation suggest involvement in continued compensation and monitoring of the new auditory feedback. The responses in the latter portion of the pitch shift, between 200-300ms, right hemisphere activity continued to increase, included peaks in right frontal (middle and inferior frontal gyrus), temporal (MTG), and parietal areas (SMG). Throughout the entire window of the pitch shift right SMG demonstrates the largest sustained power increase, indicating an involvement in coordinating the detection of and response to the pitch error. By 300ms following the pitch shift, power changes in the left hemisphere are restricted to a small region of posterior MTG, while power changes in the right hemisphere persist across SMG, PMC, and frontal regions.

4.3.5 Right parietal and premotor cortex, left posterior temporal lobe show speaking specific activity.

The motor act of vocalization is necessarily accompanied by concurrent auditory input of the acoustic consequence of one's own vocalization. By comparing the cortical responses during the beginning of compensation compared to passively listening to the same auditory input, we can identify responses that are associated with initiating vocal compensation as opposed to passively perceiving a change in pitch. Despite the widespread bilateral early activity to the unexpected pitch shift in the speaking condition, only right IPL, SMG, and PMC and left posterior MTG show greater responses in the speaking condition than in the passive listening condition (Fig 4.3, Table 4.3). The enhancement during speaking as compared to during passive listening of the response to an unexpected pitch shift, termed Speaking Perturbation Response Enhancement (SPRE) (Chang et al., 2013; Kort et al., 2014), had peak activity in right dorsal

IPL and right PMC concurrent with the onset of compensation, indicating their involvement in coordinating the motor change driving compensation. Following compensation onset, one region in the left hemisphere shows SPRE - left posterior MTG. In the right hemisphere, a new SPRE cluster was recorded in right SMG at 250ms, when the sensory consequences of the compensation should be detected.

Region	SPRE MNI peak voxel	SPRE Time-course	T-value	P-value
Left Posterior Middle Temporal Gyrus/ Occipital Lobe	-55.0 -75.0 15.0	Peak: 225-250ms Duration: 225-275ms	3.77336	4.9*10^-4
Left Middle Occipital Gyrus	-35.0 -80.0 10.0	Peak: 250-275ms Duration: 225-300ms	3.31917	0.0076
Right Premotor Cortex	55.0 -5.0 40.0	Peak: 175-200ms Duration: 100-250ms	4.08262	2.4*10^-4
Right Inferior Parietal Lobe	60.0 -25.0 30.0	Peak: 175-200ms Duration: 100-200ms	3.40243	0.0017
Right Supramarginal Gyrus	65.0 -25.0 25.0	Peak: 250-275ms Duration: 250-300ms	3.53236	0.0012
Right Occipital Lobe/ Lingual Gyrus	5.0 -90.0 -5.0	Peak: 100-125ms Duration: 100-125ms	1.90629	0.0032

Table 4.3. Table contains regions, peak voxel location, duration, peak latency, t-value and p-value of regions of significant SPRE, enhancement in response to the pitch shift during speaking compared to passive listening. 20 voxels cluster corrected, $p<0.01$.

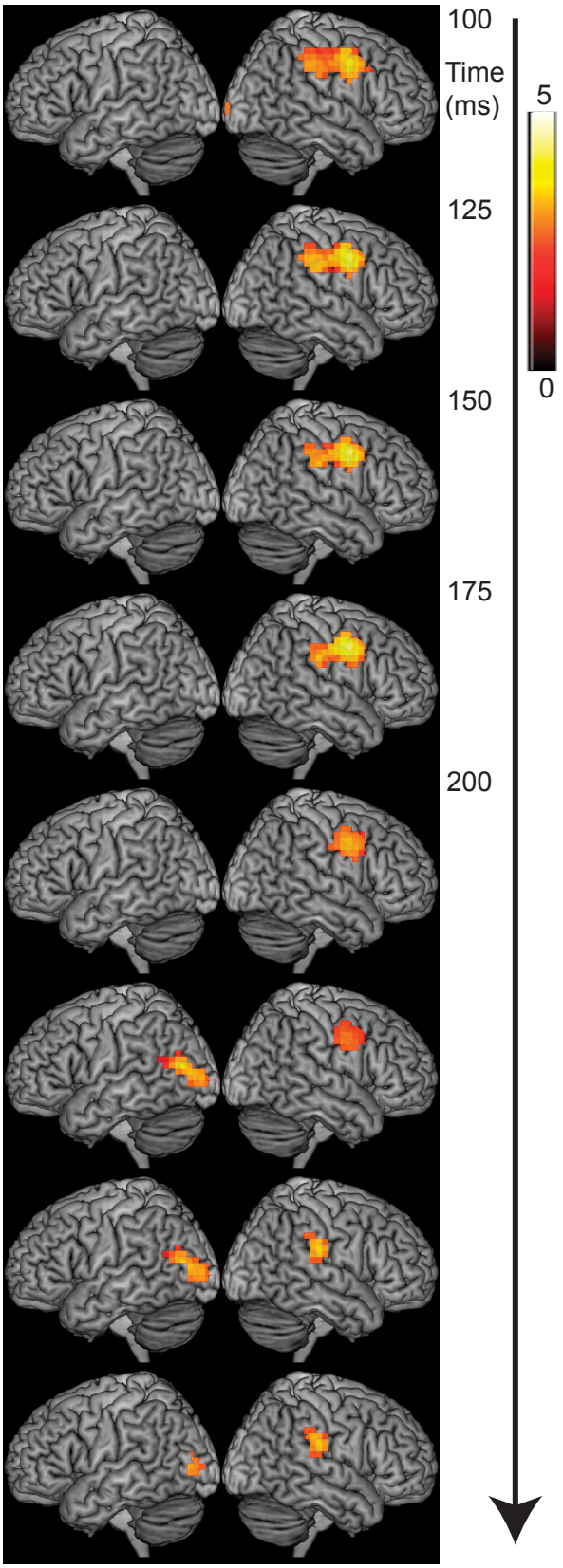

Fig 4.3. MEG cortical responses to the pitch shift that are greater in the speaking condition than the passive listening condition (SPRE). Images are 20 voxels cluster corrected, $p<0.01$.

4.3.6 Correlations between neural activity and behavioral variability of response to pitch shifts address individual subject differences.

While all subjects in this analysis compensated for the unexpected shift in pitch, the amount each subject opposed the shift varied both in the average compensation and the variance across trials. In order to address the neural underpinnings of this behavioral variability, we performed neurobehavioral correlations of HGP during speaking and individual subject's 1) mean compensation 2) variance. Neurobehavioral correlations with mean compensation revealed a network of right frontal regions, including IFG- pars triangularis, IFG- pars orbitalis and middle frontal gyrus (MFG), and left posterior temporal cortex/occipital lobe significantly positively correlated with individual subject behavioral compensation to the pitch shift (Fig 4.4, Table 4.4). Additionally, left anterior MTG showed a strong negative correlation with behavior. The neurobehavioral correlations with mean compensation occur in the first 200ms following the pitch shift. Given the timing of the correlations with mean compensation, we can infer that the initial detection and processing of the pitch shift impacts the total amount of compensation.

Neurobehavioral correlations with variance revealed two separate cortical networks (Fig 4.4, Table 4.5). Early activity- 100-125ms, in right superior temporal gyrus and right occipital lobe, was highly negatively correlated with variability in response. This indicates that stronger right hemisphere early high gamma response was highly positively correlated with increased consistency in behavior. Conversely, late (275-300ms), high gamma activity in left auditory and premotor areas and right anterior temporal and frontal regions was positively correlated with increased variability across trials. The latency of the peak of these correlations could be a result of either, or a combination, of the variability in auditory feedback as the subjects' compensate, or in the anticipation or motor change of the compensation.

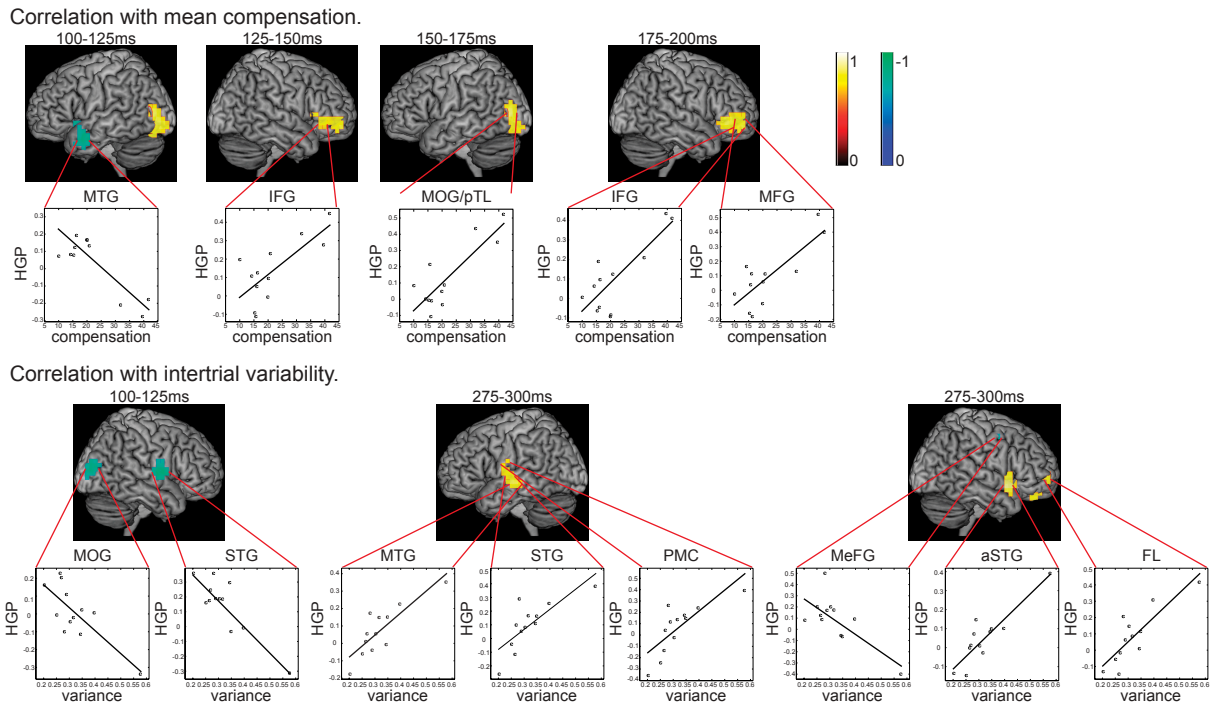

Figure 4.4. Neurobehavioral correlations across subjects with mean compensation (top panel) and variance (bottom panel). Images are 20 voxels cluster corrected, p<0.01.

Region	MNI peak voxel	Time-course	Robust R^2	P-value
Left Middle Temporal Gyrus	-55.0 0.0 -25.0	Peak: 100-125ms Duration: 100-150ms	0.6134	5.1*10^-4
Left Middle Occipital Gyrus/ Posterior Temporal Lobe	-35.0 -80.0 5.0	Peak: 150-175ms Duration: 100-250ms	0.6518	3.3*10^-4
Right Inferior Frontal Gyrus	60.0 20.0 5.0	Peak: 125-150ms Duration: 100-150ms	0.4404	0.004
Right Inferior Frontal Gyrus	35.0 35.0 0.0	Peak: 175-200ms Duration: 100-250ms	0.576	8.2*10^-4
Right Middle Frontal Gyrus	45.0 55.0 -5.0	Peak: 175-200ms Duration: 100-225ms	0.5736	0.0022

Table 4.4. Table contains regions, peak voxel location, duration, peak latency, robust r^2 and p-value of regions with significant neurobehavioral correlations with individual subject mean compensation. 20 voxels cluster corrected, p<0.01.

Region	MNI peak voxel	Time-course	Robust R^2	P-value
Left Middle Temporal Gyrus	-65.0 -15.0 -5.0	Peak: 275-300ms Duration: 275-300ms	0.6505	8.7*10^-4
Left Superior Temporal Gyrus	-65.0 -10.0 5.0	Peak: 275-300ms Duration: 275-300ms	0.5582	0.0035
Left Precentral Gyrus	-65.0 -5.0 15.0	Peak: 275-300ms Duration: 275-300ms	0.456	0.0047
Right Superior Temporal Gyrus	65.0 -5.0 5.0	Peak: 100-125ms Duration: 100-125ms	0.7432	1.9*10^-4
Right Middle Occipital Gyrus	35.0 -85.0 10.0	Peak: 100-125ms Duration: 100-125ms	0.5452	0.0010
Right Frontal Lobe (FL)	30.0 -5.0 45.0	Peak: 275-300ms Duration: 275-300ms	0.5391	0.0048
Right Anterior Superior Temporal Gyrus	45.0 10.0 -10.0	Peak: 275-300ms Duration: 275-300ms	0.7532	9.0*10^-5
Right Medial Frontal Gyrus (MeFG)	10.0 40.0 -5.0	Peak: 275-300ms Duration: 275-300ms	0.6767	9.5*10^-4

Table 4.5. Table contains regions, peak voxel location, duration, peak latency, robust r^2 and p-value of regions with significant neurobehavioral correlations with individual subject compensation variance. 20 voxels cluster corrected, $p<0.01$.

4.3.7 Inter-hemispheric effective connectivity changes in response to the external pitch shift during speaking.

In addition to increases in HGP in anatomically distinct regions, the level of functional connectivity (FC) in the high gamma band across these regions dynamically changes throughout the pitch shift. The largest and most consistent changes in FC are inter-hemispheric, specifically between left primary and secondary sensory regions and right parietal regions. In this study, functional connectivity was measured by phase-locking value (PLV) in the high gamma band (65-150Hz) during the speaking condition following the pitch shift, with speaking prior to the pitch shift as the baseline. The nodes were selected from the peaks of HGP enhancement identified in the speaking condition, and the left hemisphere analogs of the right premotor, IFG- pars triangularis, IFG- pars orbitalis and MFG peaks making a total of 17 nodes. Significant changes in FC from a steady-state vocalization baseline to the response to the pitch shift are summarized in figure 4.5 during three time windows: prior to compensation onset, during compensation onset, and following compensation onset.

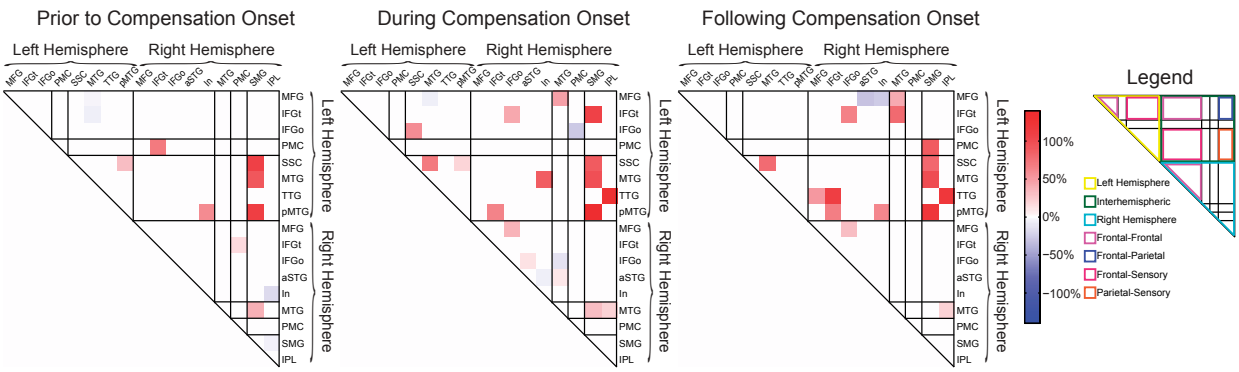

Figure 4.5. Connectivity changes in response to the pitch shift. Changes in connectivity are shown for three time windows during the pitch shift- prior to compensation onset (left), during compensation onset (middle), following compensation onset (right).

Prior to compensation onset the strongest changes in FC values were increases in inter-hemispheric links between left sensory regions, SSC, MTG and posterior MTG, and right SMG. Right SMG also shows increased FC with right MTG and slight decreased FC with right IPL. Also of particular interest during this time window, right IFG- pars triangularis shows increased FC with bilateral premotor cortex.

In the windows during compensation onset several more links show significant changes in FC values from baseline, with the most significant changes again occurring between left sensory and right parietal regions. Right SMG continues to show increased FC with several regions in both hemispheres- left SSC, MTG, posterior MTG, IFG- pars triangularis, and with right MTG. In right dorsal IPL, FC was significantly increased with left transverse temporal gyrus. Left hemisphere intrahemispheric FC changes were dominated by left SSC increased FC with several regions- left MTG, left posterior MTG and left IFG- pars orbitalis. In the right hemisphere, right MTG shows hub activity with increased FC with right SMG, right IPL, right anterior STG and decreased FC with right IFG- pars orbitalis.

Following compensation onset the strongest changes in FC values continue to be links between left sensory and right parietal regions, with the continued increased FC between right SMG and left SSC, left MTG and left pMTG. Additionally, right SMG shows increased FC with

left PMC. Left transverse temporal gyrus became a hub with increased FC with right IPL, right IFGt and right MFG. Left pMTG also shows increased inter-hemispheric FC, not only with right SMG but also right insula and right IFG- pars triangularis.

4.4 Discussion

Online control of pitch using auditory feedback is important for communication. In this study we showed the cortical mechanisms involved in online control of pitch by studying the recognition and response to a shift in pitch feedback. We demonstrate bilateral involvement in feedback control of speech. Importantly, we show left hemisphere sensory responses preceding right hemisphere frontal, parietal and premotor responses. When controlling for passive auditory perception, we find right hemisphere parietal and premotor as well as left posterior temporal responses. Further we show bilateral neurobehavioral correlations with both compensation amplitude and variance. Finally, we show large increases in inter-hemispheric communication throughout the pitch shift, the greatest between left sensory and right parietal regions. These results highlight the importance of the left hemisphere in sensory recognition and processing of the pitch shift, and the importance of the right hemisphere in higher level processing and driving the behavioral compensation.

Consistent with previous studies, we found enhanced bilateral cortical responses to an unexpected error in the pitch feedback during speaking compared with speaking with unaltered feedback (Behroozmand et al., 2011a; Chang et al., 2013; Greenlee et al., 2013; Kort et al., 2014; Parkinson et al., 2012; Tourville et al., 2008). By localizing these enhanced responses in space and time, we report the left hemisphere enhanced responses were primarily located in sensory regions- across the temporal lobe and into somatosensory cortex, and had their peak activity prior to, or in the case of posterior MTG, during compensation onset. In comparison, in

the right hemisphere, enhanced responses were more widespread including auditory, parietal, premotor and frontal regions, and with the exception of the clusters in right IFG-pars triangularis and anterior STG, occur during or following compensation onset.

The timing of the cortical activity gives insight to the mechanism of error detection and compensation, and the role of each cortical region in the circuit. In the period prior to compensation onset, left hemisphere primary and secondary sensory regions across the temporal lobe and into somatosensory cortex show enhanced responses to the pitch shift in the speaking condition, but the absence of SPRE in these regions implies very similar pattern of responses to the pitch shift also present in the listening condition. Similarly, in the right hemisphere, IFG- pars triangularis shows an early peak in the speaking condition that correlates with the amount of compensation and an absence of SPRE. Taken together, left sensory regions and right IFG appear to be involved in sensory detection and processing of the pitch shift. In the period during compensation onset, right dorsal IPL and right PMC show their peak activity both during the speaking condition and with SPRE. The timing of these peaks indicates their involvement in coordinating the motor change. Interestingly, two peaks in the left temporal lobe show their peak activity concurrent with the compensation onset with only the posterior temporal lobe showing SPRE peak later. This could indicate the posterior temporal lobe's involvement in processing both the initial pitch shift and the sensory consequence to the motor compensation. Following compensation onset the new motor plan continually increases compensation and the sensory consequence to the compensation is being detected and processed. During this time we see right SMG, right MTG and several right frontal regions show their peak activity.

Two previous studies have looked at HGP cortical responses to pitch-altered feedback using electrocorticography (ECoG) finding enhanced HGP in response to pitch-altered feedback in bilateral STG (Greenlee et al., 2013) and in left posterior temporal lobe and premotor cortex

(Chang et al., 2013). Within the left hemisphere, Chang et al. reported the latency in enhanced responses proceeded from auditory to motor regions, and in one exemplary subject showed a similar trend in the right hemisphere. These results are in accord with our study, which found left hemisphere sensory HGP enhancement preceded right hemisphere PMC enhancement. Within left STG and premotor, Chang et al. was able to find electrodes that not only showed enhanced responses to the pitch altered feedback, but that also showed correlations with individual trial compensation. Interestingly, while we did not find consist HGP enhancement across subjects in left PMC to the pitch shift, we did observe significant correlations with compensation variance in both left ventral PMC and left STG. Taken together, these results indicate left PMC in modulating the amount of compensation for an individual trial. But due to the limitation of the size of the ECoG grid, the ECoG studies could not address the role of the right hemisphere, or bilateral frontal regions.

Previous work studying the entire speech-motor network in fMRI with pitch (Parkinson et al., 2012), formant (Tourville et al., 2008) and somatosensory (Golfinopoulos et al., 2011) perturbations have provided insight into the cortical network involved in responding to sensory feedback. While the Parkinson, et al. study of pitch-altered feedback only showed bilateral STG enhancement, the restricted network shown in this finding may be due to the experimental setup. In that study, five 200ms pitch shifts were presented in each vowel utterance separated by 500-700ms of unshifted vocalization, and so results include responses to unshifted vocalization, onset of a shift and the offset of a shift. In contrast, fMRI studies of errors sustained across the whole trial have shown enhanced responses to formant altered feedback in bilateral perisylvian, right ventral SSC, motor and premotor cortices (Tourville et al., 2008), and enhanced responses to somatosensory perturbations in bilateral ventral motor, right anterior SMG, right IFG- pars triangularis and right ventral PMC (Golfinopoulos et al., 2011). Due to the temporal limitation of

fMRI, these studies were not able to address the timing of these regions, and may miss cortical regions whose responses do not persist for the duration of the shift and cannot address an error mid-utterance. Despite similarities in the cortical responses in our study to formant-altered feedback (Tourville et al., 2008), there are also several noteworthy differences. In the left hemisphere, the current findings include regions across the temporal lobe and SSC and in the right hemisphere regions include IPL, SMG, and PMC that were not seen in the Tourville et al. study. Given the timing, power, and connectivity of right SMG and right dorsal IPL in our study, the role of these regions in auditory feedback control is an important addition for any model of speech production. These aforementioned fMRI studies spurred the development of the latest version of the DIVA model, which includes a right lateralized feedback control map in right premotor cortex (Golfinopoulos et al., 2010; Tourville and Guenther, 2011). The emphasis on right premotor cortex in feedback control is in accord with our study results.

The results of the functional connectivity analysis highlight the importance of inter-hemispheric communication in recognizing and responding to pitch feedback errors. Increases in functional connectivity in response to altered auditory feedback have been demonstrated in EEG with dynamic causal modeling (Parkinson et al., 2013) showing increased connectivity between left and right STG during a pitch shift, and in fMRI with structural equation modeling showing increased connectivity between left posterior STG and right posterior STG, PMC and IFG-pars triangularis during formant-altered feedback (Tourville et al., 2008). In the current study, throughout the pitch shift the strongest increase in connectivity is between left sensory and right parietal regions. The differences in networks across the previous studies and our current study could result from the task differences, the addition of time resolution, and method of studying functional connectivity. The striking inter-hemispheric increases in connectivity following the

pitch shift highlight the importance of the coordination of the left hemisphere sensory error detection and the right hemisphere compensation and processing.

Given the right hemisphere's role in coordinating pitch production and the importance of pitch in prosody production, we postulate that the feedback control of pitch is part of the role of the right hemisphere in prosody production. This theory allows for a possible re-interpretation of prosody production deficits following stroke. Since pitch is an important element of prosody production, a subset of the deficits in prosody production could be in fact a feedback control deficit. These deficits may be completely separate from the intent to add affective or linguistic emphasis to speech. For example, patients with right hemisphere lesions leading to affective prosody production deficits (measured by pitch production) do not show improvement, and in some cases worsen performance, as verbal-articulatory demands are reduced to an utterance of the vowel /a/ (Ross et al., 1997). The failure of these right hemisphere lesion patients to appropriately modulate the pitch in the production of the vowel /a/, a similar task used in this feedback control study, shows the patients inability to appropriately control their pitch which could be due to feedback control errors. Further overlap between the pitch control network identified in the current study and the anatomical regions that disrupt prosody production includes deep white matter lesions adjacent to the corpus callosum, which disrupt prosody production (Klouda et al., 1988; Ross et al., 1997; Ross and Monnot, 2008). Given the findings in this study, the role of feedback control deficits in prosody production deficits is an important future line of study.

4.5 Conclusions

In this study, we demonstrate that both the left and the right hemisphere are involved in feedback control of speech and inter-hemispheric communication coordinates vocal feedback

control. Subjects respond to an unexpected pitch shift by changing their pitch production to oppose the change, and we found bilateral increases in high gamma power (HGP) early in response to the pitch shift change followed by predominantly right hemisphere HGP enhancement by 200ms. Speaking-specific induced HGP increases, obtained by subtracting out the cortical responses during passive listening, showed right parietal and premotor cortex (peak activity by 175ms) and left posterior temporal cortex activations (peak activity at 225ms) in driving the motor response to the shift. Neurobehavioral correlations demonstrated the involvement of right frontal regions and left posterior temporal cortex in the amplitude of the behavioral response while bilateral auditory and right frontal regions correlated with individual subjects' response variations to pitch shifts. Finally, connectivity analysis revealed large scale increases in inter-hemispheric communication that coordinates vocal feedback control of pitch.

Chapter 5.
Summary and Conclusions.

5.1 Brain networks for voice control.

The goal of this thesis was to understand the role of auditory feedback in speech production. To this end, we conducted a series of experiments to understand the real-time cortical monitoring of one's own speech production, the recognition of errors in auditory feedback, the cortical processing of these errors, and the motor (and subsequently acoustic) compensatory change.

In the first experiment we examined the cortical networks monitoring auditory feedback at both speech onset and during brief, unexpected shifts in the pitch of subjects' audio feedback during the phonation of a single vowel. We contrasted the networks between these two events. Results from this study suggest that distinct, yet overlapping networks are involved in monitoring the onset of speech and in detecting and correcting for an error during ongoing speech. Cortical monitoring at the onset of vocalization, in the form of suppressed responses during speaking compared with passive listening, was found in bilateral auditory cortex and left ventral supramarginal gyrus/posterior superior temporal sulcus (vSMG/pSTS). In contrast, cortical responses to unexpected shifts in the pitch of auditory feedback was enhanced in bilateral vSMG/pSTS, bilateral premotor cortex, right primary auditory cortex, and left higher order auditory cortex. The results of this study implicate vSMG/pSTS is involved in both monitoring the onset of speech and recognizing and responding to auditory errors during speech.

In the second study, we further investigated cortical monitoring at the onset of speech. We studied the spatial and frequency distribution of suppression during speaking compared with passive listening. We examined the effect on this suppression when the auditory feedback deviates

from the expectation, and if the cortical responses modulate with the amount the auditory feedback deviates from the expected acoustic consequence. This study confirmed previous reports, finding event related low frequency evoked suppression in bilateral temporal lobes when the perceived auditory feedback matches the expected auditory feedback. Furthermore, we demonstrated induced suppression responses in frontal, temporal, and cerebellar regions and reported that the largest extent of suppression occurs in the beta band. This suppression is reduced and eliminated as the perceived auditory feedback deviates from the expected feedback. The elimination of suppression was caused by increased activity in the same cortical regions during speaking with altered auditory feedback. Additionally, a network of regions beyond those with suppression behavior showed increased activity modulated by the amount of deviation of the perceived auditory feedback from the expected auditory feedback.

In the third study, we investigated the cortical control of voice pitch production during ongoing phonation. We studied how cortical responses to an error in pitch production evolve over time, how these responses relate to behavioral compensatory movements, and how neural connectivity in the speech motor network changes in response to an error in auditory feedback. This study showed both left and right hemisphere involvement in feedback control of speech and that inter-hemispheric communication coordinates vocal feedback control. In the first 200ms following an unexpected pitch shift, we found bilateral increases in high gamma power (HGP). By 200ms the HGP enhancement was predominantly in the right hemisphere. Speaking-specific HGP enhancement was in right parietal and premotor cortex with peak activity at 175ms following the pitch shift and in left posterior temporal cortex with peak activity at 225ms. The timing and speaking-specific enhancement of the right inferior parietal lobe and premotor cortex indicated their involvement in driving the motor compensatory response. Right frontal and left posterior temporal cortex were significantly correlated with the amplitude of the behavioral

response while bilateral auditory and right frontal regions correlated with the variance in responses a subject has to the pitch shift. Connectivity analysis revealed large scale increases in inter-hemispheric communication that coordinates vocal feedback control of pitch.

In this series of studies we have described the cortical networks and their mechanisms in monitoring auditory feedback during speech production. We have demonstrated that the cortical monitoring of the onset of a vocalization is primarily conducted with the suppression of auditory, cerebellar and frontal regions in both the left and right hemisphere. We showed this suppression is released when the perceived auditory feedback does not match the expected auditory feedback. We investigated the cortical networks involved in monitoring ongoing productions for errors in feedback, and the behavioral and cortical response when an error is perceived. We have shown that while the cortical networks monitoring the onset of speech and mid-utterance are overlapping, that these networks have distinct timing, patterns of response, and anatomical locations. Importantly, both feedback monitoring and motor control of speech at onset and mid-utterance show changes in cortical dynamics in both hemispheres. Feedback control of pitch during a mid- utterance error has corresponding increased inter-hemispheric communication. The work in this thesis provides evidence that inter-hemispheric communication is important for the feedback control of vocal pitch production.

5.2 Broader Impacts and Future Directions.

The work in this thesis has strong implications for our understanding of the neuroscience of speech. The results reported in this thesis challenge conventional models of speech production that posit left lateralization of speech production (Dronkers, 1996; Hickok et al., 2011). Instead, this thesis provides evidence that auditory error detection both at speech onset and during ongoing speech occurs in both hemispheres. The subsequent processing of this auditory error

involves inter-hemispheric communication. The right hemisphere, specifically right parietal and premotor areas, subserves the subsequent motor compensatory change. The work detailed in this thesis challenges models of speech production to address the feedback control of speech and the importance of inter-hemispheric communication to describe the neuroscience of speech production.

Understanding the neural networks involved in voice control provides insight into puzzling aspects of several neurologic and psychiatric conditions. I will outline the implications the work in this thesis has for three different conditions: stroke, Alzheimer's disease, and schizophrenia. As discussed in chapter 4, the results in this thesis allow for the re-interpretation of the cause of some prosody production deficits following stroke. Could a part of these prosody production deficits result from a feedback control deficit? While this question has not, to the knowledge of the author, been explicitly tested, there does exist some evidence that a feedback control deficit could be influencing the prosody production deficit. Specifically, some patients with right hemisphere lesions show severe deficits in pitch production when tasked with producing an intoned phonation of the single vowel /a/ (Ross et al., 1997). Further evidence for feedback control influencing prosody production deficits is the finding that inter-hemispheric communication is important in prosody production, as the ability to control prosody production is reduced following deep white matter lesions adjacent to the corpus callosum (Klouda et al., 1988; Ross et al., 1997; Ross and Monnot, 2008). Explicitly studying the effect of stroke on sensory-motor control of speech production can elucidate the underlying cause of the observed behavioral deficits.

While the majority of research into Alzheimer's disease has focused on memory related disturbances, the deterioration of motor control can have profound implications for the everyday functioning of individuals with Alzheimer's disease. In fact, decreased visual-motor control has

been reported and suggested as an early indicator of dysfunction (Tippett and Sergio, 2006; Yan et al., 2008). Given the broad networks indicated in this thesis to be involved in feedback control of speech production, and the overlap between these networks and Alzheimer's related cortical changes, it is likely that patients with Alzheimer's disease also show feedback control deficits. In a preliminary study we have shown that patients with amnesiac variant Alzheimer's disease have a disordered speech feedback control responses. A group of 9 patients with Alzheimer's disease showed significantly larger behavioral responses to an unexpected shift in the pitch of their auditory feedback during ongoing phonation of the vowel /a/ than did age matched controls. We are currently recruiting and running a large cohort of patients with Alzheimer's disease to understand the neural basis of this disorder control, how the sensory-motor feedback control of speech varies across different variants of Alzheimer's disease, how it dissociates from fronto-temporal dementia, and if sensory-motor deficits in speech production can be and early indicator for the disease progression.

Disordered pitch production can have a large impact on the quality of life and treatment of patients with several psychiatric disorders. Voice flat affect, the form of disordered pitch production in several psychiatric disorders, including schizophrenia, is characterized by decreased pitch modulation during speech. Flat affect can have profound effects on the patient's ability for social interaction, and subsequently impair a patient's treatment. The role of feedback control in causing flat affect is an important future line of study. It has been well established that people with schizophrenia show abnormal responses to the perception of their own speech (Perez et al., 2012), but the link between abnormal perception of one's own speech and the subsequent changes in speech production has yet to be established in this population.

Understanding how speech production deficits arise from errors in sensory-motor feedback control can both inform on the mechanisms of dysfunction in a disease, and can

facilitate appropriate treatment plans. A speech disorder arising from a failure to appropriately monitor auditory feedback may be able to be treated with personalized, targeted treatment utilizing intact circuits. Describing the cortical sensory-motor control of pitch production in healthy adults in this thesis is the basis with which we can examine disordered speech.

Chapter 6.

References.

Agnew, Z.K., McGettigan, C., Banks, B., Scott, S.K., 2013. Articulatory movements modulate auditory responses to speech. NeuroImage 73, 191–199.

Aliu, S.O., Houde, J.F., Nagarajan, S.S., 2009. Motor-induced Suppression of the Auditory Cortex. Journal of Cognitive Neuroscience 21, 791–802.

Andersen, R.A., Snyder, L.H., Bradley, D.C., Xing, J., 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Annual Review of Neuroscience 20, 303–330.

Bardouille, T., Picton, T., Ross, B., 2006. Correlates of eye blinking as determined by synthetic aperture magnetometry. Clinical Neurophysiology 117, 952–958.

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, 1303–1312.

Behroozmand, R., Korzyukov, O., Larson, C.R., 2011a. Effects of voice harmonic complexity on ERP responses to pitch-shifted auditory feedback. Clinical Neurophysiology 122, 2408–2417.

Behroozmand, R., Larson, C.R., 2011. Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. BMC Neurosci 12, 54.

Behroozmand, R., Liu, H., Larson, C., 2011b. Time-dependent Neural Processing of Auditory Feedback during Voice Pitch Error Detection. Journal of Cognitive Neuroscience 23, 1205–1217.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57, 289–300.

Broca, P., 1861. Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche. Bull Soc Anthropol 235–238.

Buchsbaum, B.R., Baldo, J., Okada, K., Berman, K.F., Dronkers, N., D'Esposito, M., Hickok, G., 2011. Conduction aphasia, sensory-motor integration, and phonological short-term memory â€“ An aggregate analysis of lesion and fMRI data. Brain and Language 119, 119–128.

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, 3153–3161.

Burnett, T.A., Senner, J.E., Larson, C.R., 1997. Voice F0 responses to pitch-shifted auditory feedback: a preliminary study. Journal of Voice 11, 202–211.

Chang, E.F., Niziolek, C.A., Knight, R.T., Nagarajan, S.S., Houde, J.F., 2013. Human cortical sensorimotor network underlying feedback control of vocal pitch. Proceedings of the National Academy of Sciences 110, 2653–2658.

Chang-Yit, R., Pick, H.L., Jr., Siegel, G.M., 1975. Reliability of sidetone amplification effect in vocal intensity. Journal of Communication Disorders 8, 317–324.

Chen, S.H., Liu, H., Xu, Y., Larson, C.R., 2007. Voice F[sub 0] responses to pitch-shifted voice feedback during English speech. J. Acoust. Soc. Am. 121, 1157–1163.

Cogan, G.B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., Pesaran, B., 2014. Sensory-motor transformations for speech occur bilaterally. Nature 1–7.

Cowie, R., Douglas-Cowie, E., Kerr, A.G., 1982. A study of speech deterioration in post-lingually deafened adults. J. Laryngol. Otol. 96, 101–112.
Curio, G., Neuloh, G., Numminen, J., Jousmaki, V., Hari, R., 2000. Speaking modifies voice-evoked activity in the human auditory cortex. Hum. Brain Mapp. 9, 183–191.
Dalal, S.S., Guggisberg, A.G., Edwards, E., Sekihara, K., Findlay, A.M., Canolty, R.T., Berger, M.S., Knight, R.T., Barbaro, N.M., Kirsch, H.E., Nagarajan, S.S., 2008. Five-dimensional neuroimaging: Localization of the time–frequency dynamics of cortical activity. NeuroImage 40, 1686–1700.
Dalal, S.S., Zumer, J.M., Guggisberg, A.G., Trumpis, M., Wong, D.D.E., Sekihara, K., Nagarajan, S.S., 2011. MEG/EEG Source Reconstruction, Statistical Evaluation, and Visualization with NUTMEG. Computational Intelligence and Neuroscience 2011, 1–17.
Donath, T.M., Natke, U., Kalveram, K.T., 2002. Effects of frequency-shifted auditory feedback on voice F[sub 0] contours in syllables. J. Acoust. Soc. Am. 111, 357–366.
Dronkers, N.F., 1996. A new brain region for coordinating speech articulation. Nature.
Eliades, S.J., Wang, X., 2002. Sensory-Motor Interaction in the Primate Auditory Cortex During Self-Initiated Vocalizations. Journal of Neurophysiology 89, 2194–2207.
Eliades, S.J., Wang, X., 2005. Dynamics of Auditory-Vocal Interaction in Monkey Auditory Cortex. Cerebral Cortex 15, 1510–1523.
Eliades, S.J., Wang, X., 2008. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453, 1102–1106.
Ernst, M.D., 2004. Permutation Methods: A Basis for Exact Inference. Statist. Sci. 19, 676–685.
Flinker, A., Chang, E.F., Kirsch, H.E., Barbaro, N.M., Crone, N.E., Knight, R.T., 2010. Single-Trial Speech Suppression of Auditory Cortex Activity in Humans. The Journal of Neuroscience 30, 16643–16650.
Fu, C.H.Y., Vythelingum, G.N., Brammer, M.J., Williams, S.C.R., Amaro, E., Andrew, C.M., Yaguez, L., van Haren, N.E.M., Matsumoto, K., McGuire, P.K., 2006. An fMRI Study of Verbal Self-monitoring: Neural Correlates of Auditory Verbal Feedback. Cerebral Cortex 16, 969–977.
Gage, N., Poeppel, D., Roberts, T.P.L., Hickok, G., 1998. Auditory evoked M100 reflects onset acoustics of speech sounds. Brain Research 814, 236–239.
Gelfand, J.R., Bookheimer, S.Y., 2003. Dissociating Neural Mechanisms of Temporal Sequencing and Processing Phonemes. Neuron 38, 831–842.
Golfinopoulos, E., Tourville, J.A., Bohland, J.W., Ghosh, S.S., Nieto-Castanon, A., Guenther, F.H., 2011. fMRI investigation of unexpected somatosensory feedback perturbation during speech. NeuroImage 55, 1324–1338.
Golfinopoulos, E., Tourville, J.A., Guenther, F.H., 2010. The integration of large-scale neural network modeling and functional brain imaging in speech motor control. NeuroImage 52, 862–874.
Greenlee, J.D.W., Behroozmand, R., Larson, C.R., Jackson, A.W., Chen, F., Hansen, D.R., Oya, H., Kawasaki, H., Howard, M.A., III, 2013b. Sensory-Motor Interactions for Vocal Pitch Monitoring in Non-Primary Human Auditory Cortex. PLoS ONE 8, e60783.
Greenlee, J.D.W., Jackson, A.W., Chen, F., Larson, C.R., Oya, H., Kawasaki, H., Chen, H., Howard, M.A., 2011. Human Auditory Cortical Activation during Self-Vocalization. PLoS ONE 6, e14744.
Grefkes, C., Fink, G.R., 2005. REVIEW: The functional organization of the intraparietal sulcus in humans and monkeys. Journal of anatomy 207, 3–17.
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, 133–141.
Heinks-Maldonado, T.H., Nagarajan, S.S., Houde, J.F., 2006. Magnetoencephalographic evidence for a precise forward model in speech production. NeuroReport 1–5.
Hickok, G., Houde, J., Rong, F., 2011. Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. Neuron 69, 407–422.
Hickok, G., Okada, K., Serences, J.T., 2009. Area Spt in the Human Planum Temporale Supports Sensory-Motor Integration for Speech Processing. Journal of Neurophysiology 101, 2725–2732.
Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat Rev Neurosci 8, 393–402.
Houde, J.F., Jordan, M.I., 1998. Sensorimotor Adaptation in Speech Production. Science 279, 1213–1216.
Houde, J.F., Jordan, M.I., 2002. Sensorimotor Adaptation of Speech I: Compensation and Adaptation. Journal of Speech, Language, and Hearing Research 45, 295–310.
Houde, J.F., Nagarajan, S.S., 2011. Speech production as state feedback control. Frontiers in Human Neuroscience 5, 1–14.
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, 1125–1138.
Jones, J.A., Munhall, K.G., 2002. The role of auditory feedback during phonation: studies of Mandarin tone production. Journal of Phonetics 30, 303–320.
Keough, D., Hawco, C., Jones, J.A., 2013. Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback. BMC Neurosci 14, 25.
Klouda, G.V., Robin, D.A., Graff-Radford, N.R., Cooper, W.E., 1988. The role of callosal connections in speech prosody. Brain and Language 35, 154–171.
Kort, N.S., Nagarajan, S.S., Houde, J.F., 2014. A bilateral cortical network responds to pitch perturbations in speech feedback. NeuroImage 86, 525–535.
Korzyukov, O., Karvelis, L., Behroozmand, R., Larson, C.R., 2012. ERP correlates of auditory processing during automatic correction of unexpected perturbations in voice auditory feedback. International Journal of Psychophysiology 83, 71–78.
Lachaux, J.-P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208.
Lametti, D.R., Nasir, S.M., Ostry, D.J., 2012. Sensory Preference in Speech Production Revealed by Simultaneous Alteration of Auditory and Somatosensory Feedback. Journal of Neuroscience 32, 9351–9358.
Lane, H., Tranel, B., 1971. The Lombard Sign and the Role of Hearing in Speech. Journal of Speech and Hearing Research 14, 677–709.
Lane, H., Webster, J.W., 1991. Speech deterioration in postlingually deafened adults. J. Acoust. Soc. Am. 89, 859–866.
Larson, C.R., Sun, J., Hain, T.C., 2007. Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F[sub 0] and amplitude control. J. Acoust. Soc. Am. 121, 2862–2872.
Lee, B.S., 1950. Some Effects of Side-Tone Delay. J. Acoust. Soc. Am. 22, 639–640.
Lombard, E., 1911. Le signe de l'elevation de la voix. Ann. maladies oreille larynx nez pharynx 37, 101–119.
Mormann, F., Lehnertz, K., David, P., E Elger, C., 2000. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. Physica D:

Nonlinear Phenomena 144, 358–369.
Muller-Preuss, P., Ploog, D., 1981. Inhibition of auditory cortical neurons during phonation. Brain Research 215, 61–76.
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, 577–584.
Niziolek, C.A., Nagarajan, S.S., Houde, J.F., 2013. What does motor efference copy represent? Evidence from speech production. The Journal of Neuroscience 33, 16110–16116.
Orlikoff, R.F., Baken, R.J., 1989. Fundamental frequency modulation of the human voice by the heartbeat: Preliminary results and possible mechanisms. J. Acoust. Soc. Am. 85, 888–893.
Oshino, S., Kato, A., Wakayama, A., Taniguchi, M., Hirata, M., Yoshimine, T., 2007. Magnetoencephalographic analysis of cortical oscillatory activity in patients with brain tumors: Synthetic aperture magnetometry (SAM) functional imaging of delta band activity. NeuroImage 34, 957–964.
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, 314–322.
Parkinson, A.L., Korzyukov, O., Larson, C.R., Litvak, V., Robin, D.A., 2013. Modulation of effective connectivity during vocalization with perturbed auditory feedback. Neuropsychologia VL .
Parsons, T.W., 1986. Voice and Speech Processing. Mcgraw-Hill Book Company, New York, NY.
Patel, R., Niziolek, C., Reilly, K., Guenther, F.H., 2011. Prosodic Adaptations to Pitch Perturbation in Running Speech. J Speech Lang Hear Res 54, 1051–1059.
Perez, V.B., Ford, J.M., Roach, B.J., Loewy, R.L., Stuart, B.K., Vinogradov, S., Mathalon, D.H., 2012. Auditory Cortex Responsiveness During Talking and Listening: Early Illness Schizophrenia and Patients at Clinical High-Risk for Psychosis. Schizophrenia Bulletin 38, 1216–1224.
Pfurtscheller, G., Lopes da Silva, F.H., 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical Neurophysiology 110, 1842–1857.
Poeppel, D., Phillips, C., Yellin, E., Rowley, H.A., Roberts, T.P.L., Marantz, A., 1997. Processing of vowels in supratemporal auditory cortex. Neuroscience Letters 221, 145–148.
Price, C.J., 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Annals of the New York Academy of Sciences 1191, 62–88.
Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 12, 718–724.
Robinson, S., Vrba, J., 1999. Functional neuroimaging by synthetic aperture magnetometry (SAM)., in: Recent Advances in Biomagnetism. Tokyo University Press, Sendai, pp. 302–305.
Ross, E.D., Monnot, M., 2008. Neurology of affective prosody and its functional–anatomic organization in right hemisphere. Brain and Language 104, 51–74.
Ross, E.D., Thompson, R.D., Yenkosky, J., 1997. Lateralization of Affective Prosody in Brain and the Callosal Integration of Hemispheric Language Functions. Brain and Language 56, 27–54.
Rödel, R.M.W., Olthoff, A., Tergau, F., Simonyan, K., Kraemer, D., Markus, H., Kruse, E., 2004. Human Cortical Motor Representation of the Larynx as Assessed by Transcranial Magnetic Stimulation (TMS). The Laryngoscope 114, 918–922.

Samson, S., Zatorre, R.J., 1988. Melodic and harmonic discrimination following unilateral cerebral excision. Brain and Cognition 7, 348–360.
Sekihara, K., Nagarajan, S.S., Poeppel, D., Marantz, A., 2004. Performance of an MEG Adaptive-Beamformer Source Reconstruction Technique in the Presence of Additive Low-Rank Interference. IEEE Trans. Biomed. Eng. 51, 90–99.
Shadmehr, R., Krakauer, J.W., 2008. A computational neuroanatomy for motor control. Exp Brain Res 185, 359–381.
Shum, M., Shiller, D.M., Baum, S.R., Gracco, V.L., 2011. Sensorimotor integration for speech motor learning involves the inferior parietal cortex. European Journal of Neuroscience 34, 1817–1822.
Sidtis, J.J., Volpe, B.T., 1988. Selective loss of complex-pitch or speech discrimination after unilateral lesion. Brain and Language 34, 235–245.
Singh, K.D., Barnes, G.R., Hillebrand, A., 2003. Group imaging of task-related changes in cortical synchronisation using nonparametric permutation testing. NeuroImage 19, 1589–1601.
Tippett, W.J., Sergio, L.E., 2006. Visuomotor integration is impaired in early stage Alzheimer's disease. Brain Research 1102, 92–102.
Tourville, J.A., Guenther, F.H., 2011. The DIVA model: A neural theory of speech acquisition and production. Language and Cognitive Processes 26, 952–981.
Tourville, J.A., Reilly, K.J., Guenther, F.H., 2008. Neural mechanisms underlying auditory feedback control of speech. NeuroImage 39, 1429–1443.
Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., Kuriki, S., 2007. Neural correlates of auditory feedback control in human. Neuroscience 146, 499–503.
Ventura, M., Nagarajan, S., Houde, J., 2009. Speech target modulates speaking induced suppression in auditory cortex. BMC Neurosci 10, 58.
Vrba, J., Robinson, S.E., 2001. Signal Processing in Magnetoencephalography. Methods 25, 249–271.
Yan, J.H., Rountree, S., Massman, P., Doody, R.S., Li, H., 2008. Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. Journal of Psychiatric Research 42, 1203–1212.
Yates, A.J., 1963. Delayed auditory feedback. Psychological Bulletin 60, 213–232.
Zarate, J.M., Zatorre, R.J., 2008. Experience-dependent neural substrates involved in vocal pitch regulation during singing. NeuroImage.
Zatorre, R.J., Halpern, A.R., 1993. Effect of unilateral temporal-lobe excision on perception and imagery of songs. Neuropsychologia 31, 221–232.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature  Date 6/9/2014