Auditory Cortex Activity Modulation in Response to Sensory Feedback during Audiomotor Map Learning

by

Mahima Goel

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Biomedical Imaging

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

[redacted]

Auditory Cortex Activity Modulation in Response to Sensory Feedback during Audiomotor Map Learning

by

Mahima Goel

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Biomedical Imaging

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

.... Chair

....

....

....

Committee in Charge

Acknowledgements

I would like to thank the UCSF MSBI program for giving me this opportunity of a lifetime and especially to Dr. Alastair Martin and Dr. David Saloner for always putting the program and its students first. Along with being fantastic professors, their passion for advancing science and patience with our entire cohort is nothing short of amazing.

A very special thank you to the UCSF Speech and Neuroscience Lab for allowing me to become a part of their team and for always encouraging my voice and work. In particular, I would like to thank my research mentor, Dr. Srikantan Nagarajan, for spreading his enthusiasm for science with others and playing a tremendous role in teaching me the nuances of conducting scientific experiments and writing about them. His passion for innovation and making a difference are infectious and I am grateful for being lucky enough to work with him. Additionally, the assistance of postdoctoral fellow Chang Cai during painstakingly long hours of learning new software systems along with the incredibly helpful and detailed explanations from postdoctoral fellow and former BIL lab member Megan Thompson made this thesis possible and my memories of this experience as a graduate researcher will always be associated with them.

I would also like to thank my other committee members, Dr. Ashish Raj, Dr. Janine Lupo and Dr. Heidi Kirsch, for their incredible patience during constantly moving parts of the project and for being available as mentors throughout every stage of my time as a graduate student. Not only could this project not have been possible without their guidance, but my overall time at UCSF would be incomplete without them.

Outside of UCSF's walls, I would like to thank my family for always encouraging me to do better and for believing in me during times when I did not even believe in myself. My incredible parents (Kumar Goel and Manisha Jain) and sister (Toshali Goel) are the light in my

life and have supported me through countless long nights of turmoil and anxiety with their unconditional patience and never-ending strength. Additionally, I want to thank my friends (Jessica, Prerana, Swaja and Vaibhav, in particular) for reminding me every single day that friends are the family we chose along the way. They are my biggest advocates and my pillars of strength, without whom every victory seems insignificant. This one is for you guys, thank you.

Auditory Cortex Activity Modulation in Response to Sensory Feedback during Audiomotor Map Learning

by

Mahima Goel

Abstract

Sensory feedback plays an important role in maintaining steady and fluent speech production. So far, a majority of research in speech has focused on auditory feedback while not a lot has been done on somatosensory feedback. Thus, the current study aims to further explore the effect of vocal tract somatosensory feedback on auditory cortex activity during the process of audiomotor map learning. Due to extensive evidence for a phenomenon known as motor-induced suppression (MIS), the current study hypothesizes that cortical activity will be reduced in subjects after the establishment of learning. Using an MEG touchscreen speech synthesizer set-up, subjects heard a target vowel sound and were asked to touch a location on the touchscreen that matched the sound they heard. With each trial, subjects were given feedback in the form of the sound that corresponded to the location on the screen they touched, resulting in them eventually learning to map out each vowel sound to a target location on the screen. This set-up allowed a paradigm to test for the effect of audiomotor map learning. Then, using the NUTMEG software and Champagne source localization algorithm, data from each subject was analyzed before and after learning, as well as any potential differences between auditory and motor feedback or left and right auditory cortices were noted. Findings across all 4 conditions (left auditory cortex with auditory feedback, right auditory cortex with auditory feedback, left auditory cortex with auditory and somatomotor feedback, right auditory cortex with auditory and

somatomotor feedback) in almost all the subjects found a statistically significant decrease in cortical activity following the establishment of audiomotor map learning. The current study sets up future work in which a variety of patient populations as well as different forms of feedback can be studied using a touchscreen speech synthesizing platform.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Overview

Speech production is a complex neural process which requires linguistic, phonological and motor precision. In fact, in order to achieve intelligible speech, we coordinate around 100 different muscles to produce finely tuned and well-practiced speech sounds[1]. Understanding the neural basis behind such multi-faceted processes provides critical insight into how the brain processes and integrates complex signals. Sensory feedback, in particular, plays an important role in maintaining steady and fluent speech production to the point where sensory feedback impairments can be seen in a multitude of speech disorders. So far, a majority of research in speech has focused on auditory feedback while not a lot has been done on somatosensory feedback. Thus, the current study aims to further explore the neural mechanism behind sensory feedback and its role on speech production by studying the effect of vocal tract somatosensory feedback on auditory cortex activity during the process of audiomotor map learning.

1.2 Sensory Feedback in Speech

Examples of the importance of feedback in speech are apparent all around us and there is much work in the literature to support our understanding of the feedback process. Auditory feedback, for example, has been shown to have an effect on speech in a study where experimentally delayed auditory feedback led to subjects compensating with lower speech rates[2] as well as a study in which artificially increasing the amplitude of auditory feedback led to subjects compensating with louder speech[3]. Additional support for the importance of auditory feedback can be seen in the absence of cochlear implants, where children born deaf can rarely be taught to produce more than a few utterances[4]. The importance of auditory feedback remains even after speech has been explicitly learned, showcased by studies in which post-lingual deafened adults retained basic speech intelligibility for years, but many aspects of their speech began to immediately degrade after deafness[5,6]. Similarly, visual feedback, in the form of putting children in front of a mirror and having them practice speaking, is even being used as a form of speech therapy now[7].

Somatosensory feedback, on the other hand, has not been as extensively studied. The current study, in particular, looks at a phenomenon known as motor-induced suppression (MIS) that occurs as a result of somatosensory feedback. In MIS, sensory response to stimuli is suppressed during a self-initiated task compared to the same stimuli being presented externally. For example, in a study looking at MIS, subjects were told to speak and then were later played a recording of their own voices saying the same phrase back to them. The results showed a decrease in auditory cortex activity when the subjects physically spoke (this is an example of a self-initiated task) in comparison to when the same phrase was played back to them through a recorder (this is an example of the same stimulus being presented externally)[8]. Through the paradigm of audiomotor map learning via a motor task, the current study expects to witness a decrease in audiocortical activity following sensory feedback as a result of MIS.

1.3 Clinical Relevance

Speech impairments are also noted in a wide range of neurological disorders such as Parkinson's disease-induced hypophonia, schizophrenia, autism, epilepsy and ADHD[9,10,11,12]. These disorders can lead to further downstream language and learning disorders, which are associated with developmental and social deficits[13]. In fact, around 55% of children (ages 3-5) who report having a disability also report being given speech services as part of their treatment[14].

Moreover, the specific impairment of sensory feedback is apparent in many of these speech production disorders as well. For example, in Parkinson's disease-induced hypophonia, patients are unaware of their lower than normal volume while speaking[15]. In addition, when volume or formant alterations are artificially altered in their feedback, they respond with reduced compensation to healthy controls[16]. Similarly, in stuttering, an artificial delay in auditory feedback can temporarily cause an improvement in fluency[17]. Therefore, studying the neural mechanisms of how the brain processes sensory feedback and ultimately being able to understand the complex processes of speech production and perception can have a tremendous scientific and clinical impact.

1.4 MEG Imaging

The current study utilizes MEG imaging, also known as magnetoencephalography, as the imaging modality of choice. This is due to a multitude of reasons, one of the main being that MEG imaging is a non-invasive form of functional imaging and the current study can therefore measure brain activity in a way that does not require introducing instruments into a subject's body[18]. Furthermore, unlike other non-invasive functional imaging modalities such as fMRI, MEG imaging measures direct neuronal activity by measuring microscopic changes in the magnetic field (on the order of femto-teslas to pico-teslas) that are caused by firing of neurons in the brain[19]. These changes are measured in a magnetic shielded room using magnetometers called SQUIDS (superconducting quantum intereference devices) that are bathed in liquid helium to provide low impedance by which they can amplify neuron-generated magnetic fields a few centimeters away from the sensor helmet on the participant's head[20]. In comparison, fMRI imaging utilizes a BOLD signal, which measures the blood oxygen level in various regions of the brain and assumes a correlation of brain activity from that measurement. Thus, MEG imaging offers a unique non-invasive functional imaging method that has the capability of measuring direct neuronal activity.

Moreover, MEG imaging is known to be one of the imaging modalities with the highest temporal resolutions (on the order of milliseconds) making it the ideal measuring tool for events, such as speech, that happen on rapid time scales. EEG imaging, also known an electroencephalography, also functions on similar temporal resolutions but is subject to more distortion by inhomogeneities in the scalp's conductivity and is thus not able to spatially localize signals as well as those from an MEG scanner[21,22].

Apart from being criticized for picking up a lot of noise since MEG imaging is known for measuring miniscule disturbances in the magnetic field, one of the major criticisms that MEG faces is its lack of excellence in spatial resolution (especially when compared with other functional imaging modalities, such as fMRI, that boast excellent sub-millimeter resolutions)[19]. To address this issue, the current study utilizes a novel Bayesian estimation method of source localization on the market known as the Champagne source localization algorithm[23].

1.4.1 Champagne Source Localization Algorithm

Source localization has been known to be a challenging task in MEG imaging[24]. Since the MEG instrument can pick up changes in the magnetic field on the range of femtoteslas, lots of noise in the background can often be misconstrued as signal coming from the brain. Additionally, SQUID sensors themselves can sometimes cause interference in signal acquisition as can other biological artifacts, resulting in an inherent lack of accurately being able to localize brain activity as a result of MEG signaling. In order to tackle this "inverse issue" of correctly mapping the signal back to its source brain region, however, there are a multitude of different source localization algorithms currently on the market. The current study utilizes one particular one, the Champagne algorithm, in order to estimate source activity.

Champagne has been shown to improve upon existing methods of source reconstruction in terms of accuracy, robustness and computational efficiency[25]. First, with preliminary simulated and real data, it was proven that Champagne was more robust to correlated sources and noisy data when compared with other commonly-used source localization algorithms[26]. Then, using real MEG/EEG data, Champagne was compared to three other benchmark algorithms on the market (namely, MVAB, sLORETA/dSPM and MCE)[27]. The results from this study are laid out in Figure 1.1 below. Figure 1.1A shows significant improvements in aggregate performance using the Champagne source localization algorithm with a high SNR as well as a low SNR across a range of inter-dipole distances. Furthermore, Figure 1.1B provides additional support for the robustness of Champagne by providing a visual depiction of source localization performance in cases of both a high and low SNIR of Champagne compared with three other algorithms currently being used on the market. Given the compelling evidence for accuracy and robustness in source localization compared to other algorithms that are available, the current study utilizes Champagne as the chosen method of source localization.

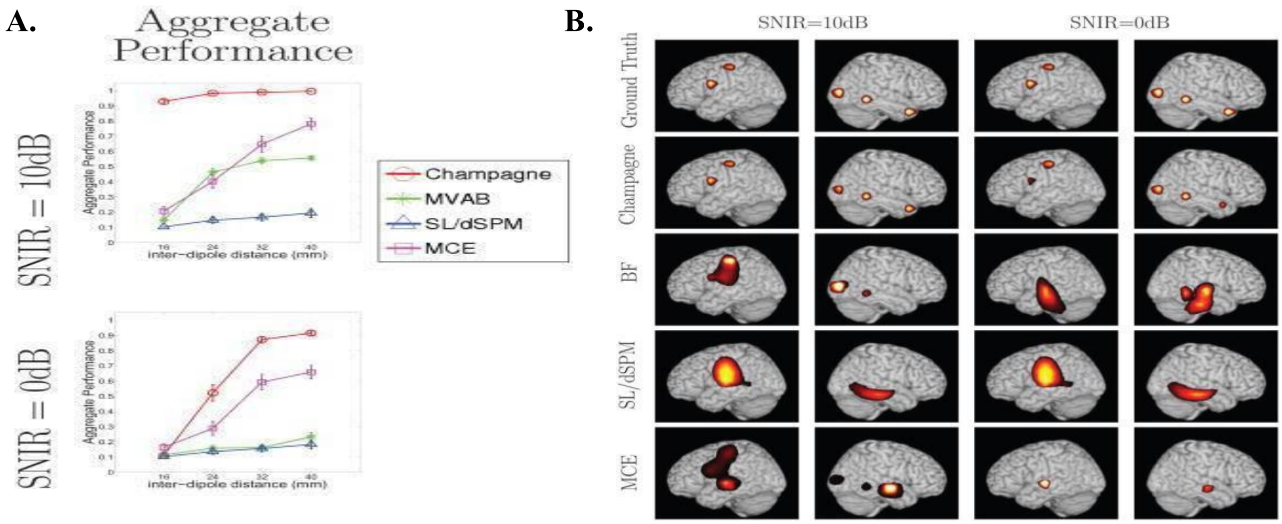

Figure 1.1A[27]: Aggregate performance of the Champagne algorithm. *Note: aggregate performance measured by the Aggregate Performance Metric (AP).* ***Figure 1.1B[27]: Single example of localization results for 5 dipoles with Champagne****. Note: ground truth shown only for comparison*

1.5 Goals & Hypothesis

Using a touchscreen speech synthesizer to replicate target sounds, a previous study was able to use a motor task with sensory feedback to establish the presence of audiomotor map learning[28]. The current study now focuses on using the said established paradigm of audiomotor map learning from the previous study to compare peak cortical activity in both the left and the right auditory cortex before and after audiomotor map learning has occurred. Furthermore, cortical activity is averaged to the onset of the target sound so as to capture the effects of solely auditory feedback as well as to the onset of the response so as to capture the effects of auditory as well as somatomotor feedback. The neural basis of MIS is thought to be based on one's expectation of what they are about to hear in comparison with what they actually hear[8]. For example, prior to learning, the auditory feedback response that they hear may not necessarily match their expectations since they have not yet developed a complete audiomotor map. On the other hand, once learning has occurred and subjects know where in the audiomotor map to expect certain sounds, the auditory feedback response will likely match their expectations. It is in this latter case that MIS has occurred and we can expect a decrease in cortical activity. Thus, the current study hypothesizes that (H1) participants will show a reduction in cortical activity after audiomotor map learning has occurred and that (H2) this trend will be seen in both the paradigm that measures solely auditory feedback as well as the paradigm that takes into account both auditory and motor feedback.

Chapter 2: Methods

2.1 Data Acquisition

2.1.1 Participants

All experimental procedures were approved by the Institutional Review Board at the University of California, San Francisco. All participants provided informed consent for their participation in said study, with a total of eighteen healthy participants (8 female, ages 18-43, mean 26.4, standard deviation 6.65). Inclusion criteria for aforementioned experiment was: 1.) participants had no prior knowledge of the mapping of screen areas to playback sounds, 2.) English was self-reported to be their first language, 3.) No known speech, language, hearing, learning or motor deficits were self-reported to be evident.

2.1.2 Equipment

A touchscreen-based speech synthesizer was attached to an MEG machine so that participant was lying supine and touchscreen was present a few meters above them at eye level. Contact with the 41.8cm by 23.6cm touchscreen (Surface Acoustic Wave Touch Panel AD-ETP-TS-LEOXXXYYZZZAA, AD Metro, Ottawa, Ontario, Canada) resulted in instant playback of a vowel sound that was dependent on the location of touch. Sound was presented through circumaural headphones to the participants and vowels were synthesized in real-time using a MATLAB program adapted from the Rabiner and Schafer Vowel Synthesis toolbox[29]. Axes of the touchscreen were associated with continuous F2 (800Hz-2500Hz) and F1 (100-900Hz) formant frequencies to span even the most conservative ranges for vowel discrimination in the literature[30] with selective fixed values of F0 corresponding to 100Hz, F3 corresponding to

2500Hz and F4 corresponding to 4000Hz. All locations corresponded to a clear, intelligible adult male speaker.

2.1.3. Procedures

Experimenters investigated whether participants could reproduce specific vowel sounds by touching a target location on the touchscreen. Over the course of 4 trial blocks of 120 trials each, participants were asked to repeat one of three randomized auditory vowel sounds (consisting of /a/, /e/, and /i/) by touching a target location on the screen that matched the given sound. For each trial, participants reached from the same starting position (lower right-hand corner of the screen, within a range of 52.25mm x 59mm) to a point on the screen. The /a/, /e/, and /i/ targets were 191.28mm, 271.60mm and 362.06mm away from the bottom-right corner of the touchscreen, respectively. A layout of the touchscreen can be visualized in Figure 2.1. Once they made contact, participants were met with feedback that corresponded to the auditory vowel sound that resulted from a combination of formants that mapped the coordinates of the location they contacted on the screen. Participants were further instructed to close their eyes for the duration of the experiment to eliminate interference from visual feedback.

The vowel targets were presented to participants at random with equal frequencies of occurrence and a set loudness level of 60dB and duration of 600ms. These particular three target vowels were selected due to their frequent usage in speech experiments and that, together, they largely span the space of the touchscreen. Similarly, formant values for target locations were selected to reflect mean formant production values for adult male speakers[31,32] subject to minor adjustment for synthesis sound quality and clarity in accordance with KTH Royal Institute of Technology's text-to-speech vowel synthesis[33,34].

Once participants made contact with the starting position, a 300ms "continue" sound consisting of a pure 1000Hz tone played. This frequency was specifically chosen in an effort to distinguish this tone from the rest of the synthesized vowel sounds as much as possible so as to not integrate into the audiomotor map. Following this tone was a 600-900ms randomly jittered silent delay period, which was subsequently followed by auditory presentation of the next target vowel sound. After participants attempted to map the location of the vowel sound on the touchscreen, the auditory feedback they received was also displayed with the same parameters of the target sound (60dB loudness and duration of 600ms).

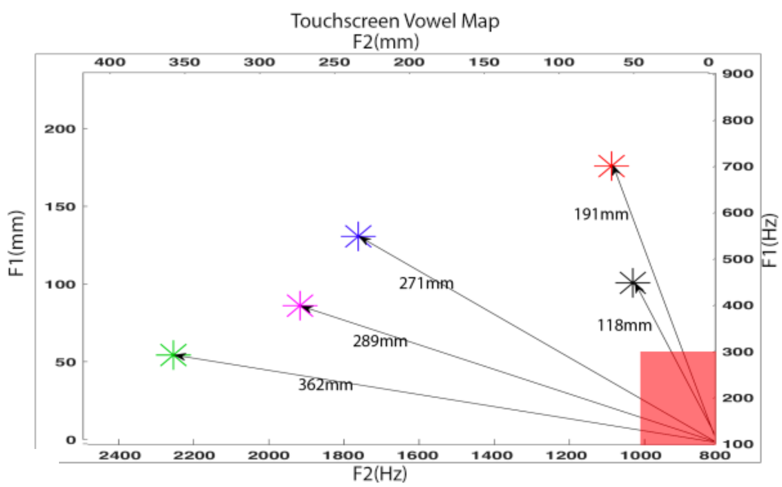

Figure 2.1[28]: Visual representation of touchscreen layout. *Note: both dimensions in Hz as well as mm are included. Starting position is depicted in red (bottom right hand corner) and trained targets are indicated with a colored * (red for /a/, blue for /e/, and green for /i/)*

2.1.4 Presence of Learning Established

Results from participants across 480 trials are shown below in Figure 2.2. Accuracy of the mapped target location was measured by averaging the Euclidean distance of the point of contact from the actual point of the target sound. As can be seen by Figure 2.2A, average accuracy across all subjects drastically and rapidly improved even within the first 30 trials (p=0.5697e-14) for all trained targets (p=1.381e-6). For /a/, accuracy improved from 187.66mm (+/-21.5) to 42.48mm (+/-9.07), whereas for /e/, accuracy improved from 102.23mm (+/-11.5) to

56.11mm (+/-10.26) and for /i/, accuracy improved from 161.61mm (+/-18.45) to 61.22mm (+/-9.96). Furthermore, consistency (measured by averaging the Euclidean distance between responses in Figure 2.2B) also drastically and rapidly improved within the first 30 trials (p=1.261e-23). Collectively, these results indicate the presence of audiomotor map learning and set the parameters for the current study to study the effects of sensory feedback on cortical activity during this process of audiomotor map learning.

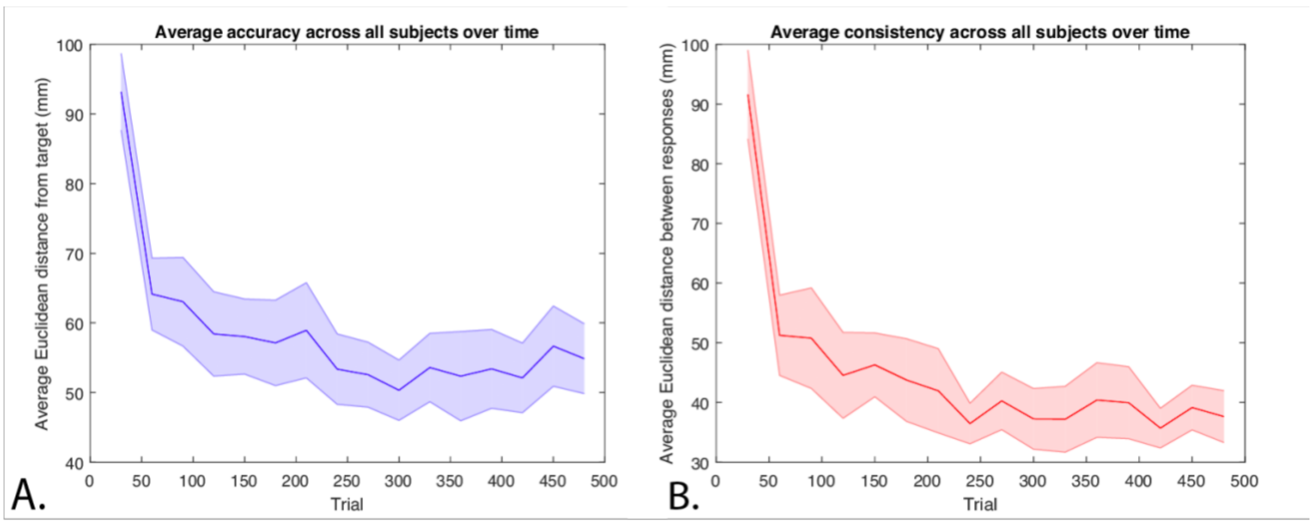

Figure 2.2A[28]: Response accuracy in identifying mapped target vowel sound locations. *Note: improvement measured by a decrease in average Euclidean distance from target over the course of 480 trials.* ***Figure 2.2B[28]: Response consistency in identifying mapped target vowel sound locations.*** *Note: improvement measured by a decrease in average Euclidean distance between responses over the course of 480 trials.*

2.2 Data Analysis

2.2.1 Overview of NUTMEG Pipeline

NUTMEG (Neurodynamic Utility Toolbox for Magnetoencephalo- and Electroencephalo- Graphy) is the chosen source analysis toolbox used to conduct data analysis in the current study. Primarily written in MATLAB and interoperable with a variety of other software, some other advantages of using NUTMEG include: 1. being able to utilize an inverse algorithm of choice, 2. intuitive viewing and navigation of results, and 3. several methods of source space functional connectivity analysis[35]. The NUTMEG workflow contains multiple steps, illustrated in Figure 2.3. The first step involves loading in the MEG data as well as MRI and coregistration information, while the next step focuses on preprocessing the data from the MEG sensor map and applying that onto the head model to calculate spatial filter weights. Finally, the third step is to use the spatial filter weights to reconstruct the source of the signal (in this case, we have utilized Champagne as the chosen source reconstruction algorithm) and then run statistics on the resulting data.

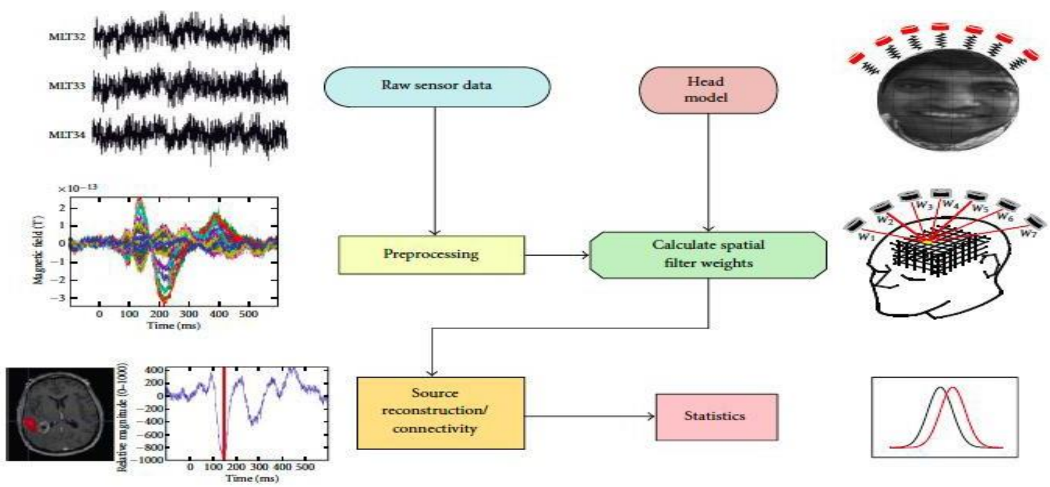

Figure 2.3[35]: NUTMEG workflow summarized in a visual flowchart

2.2.2 Raw Sensor Data

From the previous study that established audiomotor map learning, raw MEG data across 120 trials was averaged to create 4 blocks totaling 480 trials in order to phase-lock cortical activity across all trials in each block. Data from block 1 and block 4 (prior to learning and after learning, respectively) was analyzed in the current study across 17 subjects. Due to an undetected contraindication to MRI, one of the data sets from this previous study is excluded from data analysis in the current study. Each block was averaged in two forms using the DataEditor software: 1. averaged to the onset of the target vowel sound playing represented solely auditory feedback, whereas 2. averaged to the onset of the response sound being played back represented auditory as well as components of somatomotor feedback. Therefore, raw sensor data resulted in 4 files of 120 trials each averaged for 17 different subjects: 1. block 1 (prior to learning) averaged to target vowel sound, 2. block 1 (prior to learning) averaged to response sound playback, 3. block 4 (after learning has occurred) averaged to target vowel sound, and 4. block 4 (after learning has occurred) averaged to response sound playback. For the purposes of simplicity, these 4 data sets will be referred to as the 4 "conditions" from here on out. In this way, the effects of auditory feedback as well as auditory and somatomotor feedback combined can be interpreted during the process of audiomotor map learning.

2.2.3 Preprocessing

After loading the raw data into NUTMEG, a specific time window of interest as well as specific sensors to be picked for further processing can be specified. Furthermore, baseline removal, filtering and automated artifact rejection can also be applied. In the current study, the time window was the entirety of the trial (all 1800 time points were kept) and all sensors were selected to be processed (totaling to approximately 275 channels). Moreover, the current study utilized a bandpass filter of 1-100Hz in order to remove additional noise.

2.2.4 Head Model

NUTMEG includes a default multisphere model for MEG head models, but also includes an option to import individual subject's structural MRIs. In the current study, each subject's structural MRI was spatially normalized and the corresponding fiducials were manually imported using the SPM8 software. Then, using the Coregistration Tool Graphical User Interface in NUTMEG, the spatially normalized MRIs and fiducials were used to create a lead field model that was then used as a model for source estimation. Note: All lead fields were specified at an 8mm resolution.

2.2.5 Spatial Filter Weights

Applying the Champagne source localization algorithm on the lead field model, spatial filter weights were extracted. Champagne combines SEFA modeling of background noise with sparse Bayesian inference of source activity in all voxels simultaneously using fast, robust update rules with guaranteed convergence under many realistic conditions, bearing some similarities to the SAKETINI model with the difference being that SAKETINI considers each voxel sequentially while Champagne considers all voxels simultaneously[35]. This step of the data analysis was conducted with the help of postdoctoral fellow, Chang Cai.

2.2.6 Source Reconstruction

Multiplying the processed MEG signal with the calculated spatial filter weights resulted in 8 individual time series for each subject. This process was done manually, where the MEG signal voxels were first plotted onto a 3D graph and then subsequently split into the left auditory cortex (LAC) and the right auditory cortex (RAC) by dividing the plotted brain voxels in half. Then, the power of each individual voxel within the given region (either LAC or RAC) is calculated and the voxel with the maximum power is found. The power of said voxel is defined as peak cortical activity and is plotted over the course of the duration of the trial (1800 time points). This resulting plot is two time series of the peak cortical activity plotted against time points for block 1 (first 120 trials) and block 4 (last 120 trials). Each of the four conditions resulted in two time series of block 1 and block 4, ultimately resulting in 8 time series total from each subject: LAC before learning (averaged to onset of target vowel sound), LAC after learning (averaged to onset of target vowel sound), RAC before learning (averaged to onset of target vowel sound), RAC after learning (averaged to onset of target vowel sound), LAC before learning (averaged to onset of response sound playback), LAC after learning (averaged to onset of response sound playback), RAC before learning (averaged to onset of response sound playback), RAC after learning (averaged to onset of response sound playback). A visual representation of the multi-step process of obtaining two time series for one of the conditions in one subject is shown in Figure 2.4 below.

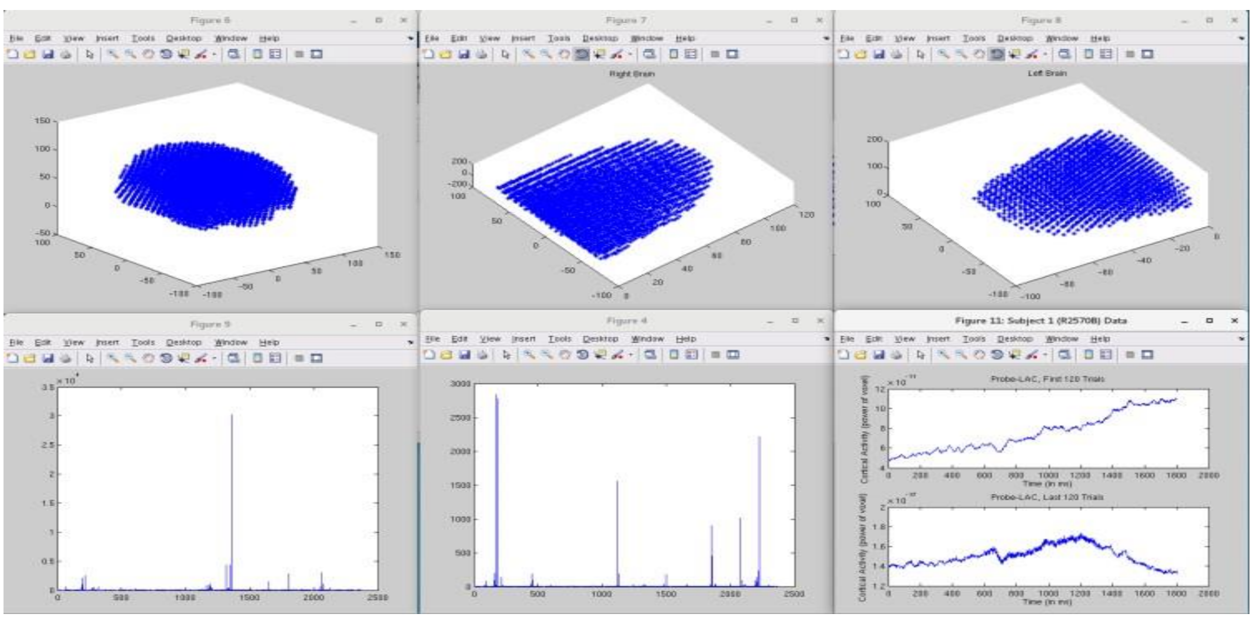

Figure 2.4: Multi-step process of obtaining two time series of peak cortical activity. *1. Plot 3D graph of MEG signal voxels (top left), 2,3. Split into LAC and RAC (top middle, top right), 4,5. Calculate voxel power and identify voxel with maximum power, 6. Plot peak cortical activity (measured by maximum voxel power) against time (in milliseconds) to obtain time series*

2.2.7 Statistics

From each time series, peak cortical activity was measured by extracting the maximum point of cortical activity across all 1800 time points. The average of these maximum values across all 17 subjects was considered the average of peak cortical activity. Furthermore, differences between peak cortical activity before and after the establishment of learning (prior to learning minus after learning) were calculated as a measure of the degree of decrease in cortical activity. A paired t-test was utilized (specifically, the Wilcoxon Ranked Sign test to account for nonparametric data) in all 4 conditions to compare the statistical significance of cortical activity before and after learning as well as to compare the statistical significance of any possible differing trends between the LAC and RAC and between auditory feedback and motor feedback. Note: When measuring for cortical activity before and after learning, peak cortical activity was used in contrast with measuring for differences between auditory and motor feedback (in which the difference between activity prior to learning and activity after learning was used). The threshold for statistical significance was set at $p \leq 0.01$ (printed p</=0.01) and all values were measured to 5 significant figures.

Chapter 3: Results

3.1 Overview of Trends

In accordance with our first hypothesis (H1), subjects showed a decrease in cortical activity following the presence of audiomotor map learning. Furthermore, this trend—in accordance with our second hypothesis (H2)—held true during both the paradigm that measured solely for auditory feedback as well as the paradigm that measured for both auditory as well as motor feedback. The decrease in cortical activity between the auditory feedback paradigm was noted to be less than the decrease in cortical activity between the auditory and motor feedback paradigm with a medium effect size, but this effect was not statistically significant. Similarly, differences between reduction in the LAC in comparison to reduction in the RAC were also noted to not be of statistical significance.

3.2 Before vs. After Learning

A decrease in cortical activity following the presence of audiomotor map learning can be seen across all 4 paradigms/conditions in Table 3.1. Peak cortical activity was measured as the maximum point of cortical activity through the duration of individual time series (8 total from each subject). Moreover, this trend is statistically significant (Note: A limit of p</=0.05 was used to measure statistical significance) across each condition: LAC activity averaged to onset of target vowel sound decreased after learning with a medium effect size of 0.368 (p=0.002000), LAC activity averaged to onset of response sound playback decreased after learning with a medium effect size of 0.348 (p=0.0003740), RAC activity averaged to onset of target vowel sound decreased after learning with a medium effect size of 0.426 (p=.0058480), and RAC activity averaged to onset of response sound playback decreased after learning with a small effect size of 0.265 (p=0.0200000). Each condition had 3 subjects in which this trend did not

hold true (*), but it was noted that each of these outliers differed by no more than 3 magnitudes of each other (in comparison to the range of differing magnitudes from 1 to 10 seen in subjects where cortical activity did in fact decrease after the establishment of audiomotor map learning). Figure 3.1 compares the average peak cortical activity across all the subjects before and after learning with the standard error bars being derived from the standard deviation of the means.

Peak Cortical Activity in Time Series for 8 Time Series

Left Auditory Cortex Averaged to Onset of Target Vowel Sound (Auditory Feedback)				Left Auditory Cortex Averaged to Onset of Response Sound Playback (Auditory and Motor Feedback)			
Subject	Prior to Learning	After Learning	Difference (Prior-After Learning)	Subject	Prior to Learning	After Learning	Difference (Prior-After Learning)
S1	1.1060E-10	1.7362E-17	1.1060E-10	S1	2.6000E-03	2.0036E-17	2.6000E-03
S2	4.7756E-05	3.4769E-07	4.7408E-05	S2*	6.3316E-05	1.1876E-04	-4.8550E-05
S3*	1.0338E-09	2.1254E-09	-1.3471E-04	S3	8.7922E+02	1.6369E-08	8.7922E+02
S4	4.2400E-02	1.2700E-02	2.9700E-02	S4	2.1900E-02	2.0740E-08	2.1900E-02
S5	1.3470E+01	3.8228E-01	1.3088E+01	S5	5.4368E+01	2.4520E-01	5.4123E+01
S6	4.7910E-01	3.0900E-02	4.4820E-01	S6	6.4572E+00	1.3930E-01	6.3179E+00
S7*	1.6730E-09	4.4335E-08	-1.2810E-03	S7	2.3953E+00	1.3709E-09	2.3953E+00
S8	9.3221E-12	2.7089E-14	9.2950E-12	S8	1.5708E-10	6.4380E-13	1.5644E-10
S9*	2.1934E-11	1.1408E-10	-1.5158E-05	S9*	3.4423E-11	1.5646E-10	-1.2157E-10
S10	3.1932E-13	2.0469E-20	3.1932E-13	S10	8.2183E+01	2.7607E-06	8.2183E+01
S11	5.5900E-02	5.0280E-08	5.5900E-02	S11*	2.0391E-07	2.2195E-04	-2.2179E-04
S12	1.7864E+02	1.4660E-07	1.7864E+02	S12	2.5711E+02	6.3559E-06	2.5711E+02
S13	1.2000E-02	7.2674E-04	1.1273E-02	S13	1.4300E-02	1.0000E-03	1.3300E-02
S14	1.4453E+02	1.6990E-01	1.4436E+02	S14	2.9850E-01	6.1627E-05	2.9844E-01
S15	1.9037E-07	6.9575E-08	1.2080E-07	S15	2.3499E-07	8.1079E-08	1.5391E-07
S16	8.6000E-03	2.2942E-08	8.6000E-03	S16	2.3455E+00	1.1600E-02	2.3339E+00
S17	9.4800E-02	7.7400E-02	1.7400E-02	S17	3.0250E-01	1.1570E-01	1.8680E-01
Average	1.9843E+01	3.9641E-02	1.9803E+01	Average	7.5572E+01	3.0189E-02	7.5542E+01

Right Auditory Cortex Averaged to Onset of Target Vowel Sound (Auditory Feedback)				Right Auditory Cortex Averaged to Onset of Response Sound Playback (Auditory and Motor Feedback)			
Subject	Prior to Learning	After Learning	Difference (Prior-After Learning)	Subject	Prior to Learning	After Learning	Difference (Prior-After Learning)
S1	2.2528E-08	2.7574E-16	2.2528E-08	S1	4.1937E-08	2.9742E-16	4.1937E-08
S2*	2.2100E-02	4.5290E-01	-4.3080E-01	S2*	2.6230E-01	1.0433E+00	-7.8100E-01
S3*	6.3289E-04	6.6000E-03	-5.9670E-03	S3*	2.3000E-03	8.7940E-01	-8.7700E-01
S4	8.9000E-03	3.2000E-03	5.7000E-03	S4	1.3600E-02	1.6504E-15	1.3600E-02
S5	1.7000E-03	2.9746E-15	1.7000E-03	S5	2.4000E-03	6.3793E-15	2.4000E-03
S6	2.8210E+01	2.4439E-09	2.8210E+01	S6	5.4008E+02	6.3608E-08	5.4008E+02
S7	8.5310E-01	3.9905E-07	8.5310E-01	S7	5.0000E-03	2.9526E-03	2.0474E-03
S8*	2.8396E-11	1.0564E-09	-1.0274E-09	S8	5.3766E-04	1.4217E-10	5.3766E-04
S9	3.9209E+00	6.4776E-04	4.7408E-05	S9	2.7311E-05	3.7454E-09	2.7307E-05
S10	5.3214E+00	2.2373E-05	5.3214E+00	S10	1.1765E+01	4.3876E-05	1.1764E+01
S11	4.2100E-02	1.3806E-06	4.2099E-02	S11	2.0580E-01	2.4023E-05	2.0578E-01
S12	2.0000E-03	6.8170E-12	2.0000E-03	S12	3.6254E+01	1.5474E-06	3.6254E+01
S13	2.1000E-03	8.2131E-10	2.1000E-03	S13	9.0000E-03	4.3547E-09	9.0000E-03
S14	3.7513E-05	1.4249E-07	3.7371E-05	S14*	1.6870E-06	2.2330E-05	-2.0643E-05
S15	9.5100E-01	1.5519E-05	9.5098E-01	S15	2.4081E+00	2.7056E-05	2.4081E+00
S16	5.0200E-02	3.6230E-12	5.0200E-02	S16	7.0700E-02	1.1196E-13	7.0700E-02
S17	1.6000E+01	8.4288E-11	1.6000E+01	S17	8.9425E-05	3.4663E-10	8.9425E-05
Average	3.2580E+00	2.7258E-02	3.0002E+00	Average	3.4770E+01	1.1328E-01	3.4656E+01

Table 3.1: Peak cortical activity for 8 time series. *LAC averaged to onset of target (top left), LAC averaged to onset of response (top right), RAC averaged to onset of target (bottom left), RAC averaged to onset of response (bottom right). Difference (fourth column in each section) was measured by subtracting the peak cortical activity value prior to learning from the peak cortical activity after the establishment of learning. Note: Negative values (*) were included in the calculations for average cortical activity*

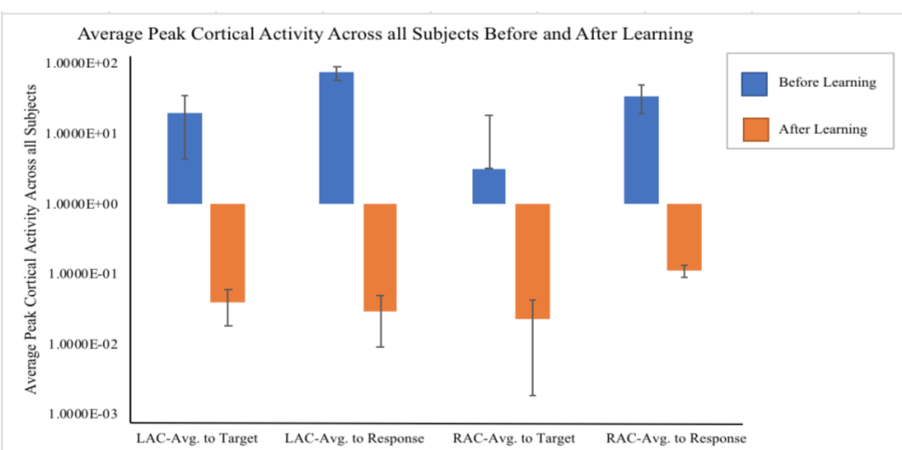

Figure 3.1: Average peak cortical activity across all subjects before and after learning. *Note: A logarithmic scale is used to account for magnitudes of change in cortical activity; before learning is shown by blue and after learning is indicated by orange*

3.4 Auditory vs. Motor Feedback

In order to evaluate any potential differences between the conditions averaged to the onset of the target vowel sound (a paradigm for measuring solely auditory feedback) with the conditions averaged to the onset of the response sound playback (a paradigm for measuring auditory as well as motor feedback), differences in activity (after learning minus prior to learning peak cortical activity, column 4 from Table 3.1) were compared. Across both conditions, no differences were found to be statistically significant: LAC activity averaged to onset of response decreased when compared to LAC activity averaged to onset of target with a medium effect size of 0.56 ($p=0.0770992$) while RAC activity averaged to onset of response decreased when compared to RAC activity averaged to onset of target with a small effect size of 0.16 ($p=0.178180$). Note: In these conditions, prior to learning and after learning blocks were combined together to measure cortical activity averaged to a specific onset and were not compared against each other.

Similarly, in order to analyze any potential differences with reduced cortical activity between the LAC and the RAC, differences in activity (column 4 from Table 3.1) between each of the 4 conditions in the LAC were compared to their corresponding condition in the RAC. Across all 4 conditions, no differences were found to be statistically significant: RAC activity averaged to onset of target vowel sound decreased when compared to LAC activity averaged to onset of target vowel sound with a medium effect of 0.46 ($p=0.510170$), LAC activity averaged to onset of response sound playback decreased when compared to RAC activity averaged to onset of response sound playback with a medium effect of 0.41 ($p=0.650307$). Note: In these conditions as well, prior to learning and after learning blocks were combined together to measure cortical activity averaged to a specific onset and were not compared against each other.

Chapter 4: Discussion

4.1 Interpretation of Trends

In accordance with our first hypothesis (H1), almost all the subjects showed a statistically significant decrease in cortical activity following the presence of audiomotor map learning. This finding hints at the presence of MIS and provides further support for the neurological mechanism of a reduction in brain activity following a subject's expectations matching with the presented auditory stimulus. Furthermore, this trend—in accordance with our second hypothesis (H2)—is robust enough to remain true when the current study analyzes both paradigms with auditory feedback as well as a combination of auditory and motor feedback. This result gives us insight from a neurological perspective on how different types of feedback are processed in speech, specifically suggesting that each form of feedback (in this case, auditory or motor) are not independent of each other since when both are present they do not have an additive effect leading to a higher reduction in cortical activity. Rather, this result provides support for a combined feedback processing mechanism in which the auditory cortex and motor cortex collectively receive and analyze speech signals. Previous work has looked at the potential interactions between the auditory and motor cortices during speech[36,37], but few have looked at the specific differences in cortical activity between auditory and motor feedback. Thus, these results have the potential to provide novel and important evidence in support for a coupling mechanism between the auditory and motor cortex during speech processing.

Moreover, seeing a smaller difference in average cortical activity after the establishment of learning (by a magnitude of 2) specifically in the LAC across all the subjects inspired the current study to analyze any potential differences in cortical activity between the two lobes of the brain. While the effect size between the cortical activity difference in the two lobes was deemed statistically insignificant, the fact that the LAC showed a two-fold larger value of cortical activity even after the establishment of learning than the RAC did provides further support for the well-known phenomenon that the left hemisphere auditory areas of the brain are more active during the process of speech than the right[38,39,40,41]. Previous studies on brain laterality have shown auditory areas in the left hemisphere of the brain to be primarily involved in the acoustical analysis of speech signals[42], phonological mapping of speech signals to conceptual representations in the brain[43,44,45], as well as the semantic processing of words[46,47]. Thus, by showing a more active LAC the current study hints at the dominant trend of the left brain continuing during the process of auditory feedback as well.

4.2 Limitations & Future Directions

As is true of all research, the current study also had a series of limitations. Amongst the most prominent was having a small sample size of 17 subjects, resulting in the inherent limitation of weak to medium statistical power. Furthermore, since MEG data is known to be noisy due to hyper-sensitivity on SQUID sensors, it is difficult to be able to completely shield the data from noise. Even within our sample, despite multiple rounds of filtering and cleaning the signal through NUTMEG, some data sets contained choppy portions of noise. This noise can be attributed to improper shielding in the magnetically shielded MEG room, interference from the SQUID sensors themselves causing disruptions, or other external factors such as excessive movement or potential feelings of nervousness that can cause physiological responses and alter brain activity.

Learning from these limitations, the current study hopes that future work will be done on increasing the filtering ability of NUTMEG as well as possibly developing new algorithms for making MEG data less noisy than the status quo. Studying other forms of feedback using a similar touchscreen platform set-up for speech is another area for future work that has tremendous potential (i.e. analyzing the effects of musical training, analyzing the effects of visual feedback by allowing subjects to keep their eyes open, analyzing the effects of being able to speak vowels aloud while touching their location on the screen, etc.) The potential for utilizing this touchscreen platform set-up for speech can further be extended by studying the effects of auditory and somatomotor feedback across a variety of different populations with speech impairments (i.e. patients with speech disorders such as spasmodic dysphonia, Parkinson's-induced hypophonia, apraxia, dysarthria, and aphasia or patients with disorders that can result in speech defects such as ADHD, Huntington's disease, dementia, ALS, and autism). Lastly, the number of applications for the Champagne source localization algorithm are practically boundless, ranging from analyzing MEG data on patients to controls across a plethora of different experimental set-ups and paradigms. The current study hopes to inspire and propel more studies in the future that will utilize the Champagne algorithm to better understand the neural mechanisms behind the complex and multifaceted processes of speech production and perception.

Chapter 5: Conclusion

From a previous experiment, the presence of audiomotor map learning was established. In this touchscreen speech synthesizing set-up, subjects were asked to match a location on the screen with a target vowel sound that they heard. Subjects were also given feedback as to the sound that corresponded with the location on the screen that they chose, allowing them to learn over the course of many trials how to map each vowel sound to a target location on the screen. Using this experiment as a paradigm for audiomotor map learning, the current study analyzed differences in cortical activity before and after the presence of audiomotor map learning across all subjects. The NUTMEG software pipeline was utilized to filter, sort and categorize the MEG data, whereas the Champagne source localization algorithm allowed accuracy in pinpointing the signal to either the LAC or RAC. Data was split into before and after learning, as well as averaged to the target vowel sound (giving us a paradigm to study solely auditory feedback) and averaged to the response sound playback (giving us a paradigm to study auditory as well as motor feedback). Data was further split into the LAC and RAC, allowing the current study to analyze brain laterality along with the effects of specific forms of feedback during learning. Across all 4 conditions, a statistically significant decrease in cortical activity was noted after the establishment of audiomotor map learning. This suggests the presence of MIS and provides support for our first hypothesis (H1). Furthermore, no statistically significant differences in cortical activity were found between auditory versus somatomotor feedback, hinting at a coupling interaction between the auditory and motor cortices in processing speech. Although the differences were not significant, the fact that both conditions (auditory feedback as well as a combination of auditory and somatomotor feedback) resulted in decreased cortical activity ultimately supports our second hypothesis (H2). Moreover, higher activity after learning in the LAC (by a magnitude of 2) suggests support for previous literature claiming that the left hemisphere auditory areas of the brain are more active during speech. Now, future work can focus on further exploring the touchscreen speech synthesizer paradigm as well as NUTMEG and the Champagne algorithm to better understand the neural mechanisms behind the complex processes of speech production and perception.

References

[1] Darley, Frederic L., Arnold Elvin Aronson, and Joe Robert Brown. *Motor speech disorders*. Saunders, 1975.

[2] Lee, Bernard S. "Effects of delayed speech feedback." *The Journal of the Acoustical Society of America* 22.6 (1950): 824-826.

[3] Behroozmand, Roozbeh, and Charles R. Larson. "Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback." *BMC neuroscience* 12.1 (2011): 54

[4] Kishon-Rabin, Liat, et al. "The effect of partially restored hearing on speech production of postlingually deafened adults with multichannel cochlear implants." *The Journal of the Acoustical Society of America* 106.5 (1999): 2843-2857.

[5] Cowie, Roddy, and Ellen Douglas-Cowie. *Postlingually acquired deafness: speech deterioration and the wider consequences*. Vol. 62. Walter de Gruyter, 1992.

[6] Waldstein, Robin S. "Effects of postlingual deafness on speech production: implications for the role of auditory feedback." *The Journal of the Acoustical Society of America*88.5 (1990): 2099-2114.

[7] Williams, A. Lynn, Sharynne McLeod, and Rebecca J. McCauley. *Interventions for Speech Sound Disorders in Children*. Brookes Publishing Company. PO Box 10624, Baltimore, MD 21285, 2010.

[8] Houde, John F., et al. "Modulation of the auditory cortex during speech: an MEG study." *Journal of cognitive neuroscience*14.8 (2002): 1125-1138.

[9] Scott, S., F. I. Caird, and B. O. Williams. "Evidence for an apparent sensory speech disorder in Parkinson's disease." *Journal of Neurology, Neurosurgery & Psychiatry* 47.8 (1984): 840-843.

[10] Thompson, Megan. *Speech and Nonspeech Production I the Absence of the Vocal Tract*. Diss. University of California, San Francisco, 2018.

[11] Goldberg, Terry E., et al. "Unnatural practices, unspeakable actions: a study of delayed auditory feedback in schizophrenia." *The American journal of psychiatry* 154.6 (1997): 858.

[12] Croft, L. J., et al. "To speak, or not to speak? The feasibility of imaging overt speech in children with epilepsy." *Epilepsy research* 107.1-2 (2013): 195-199.

[13] Yorkston, Kathryn M., David R. Beukelman, and Kathleen R. Bell. *Management of motor speech disorders in children and adults*. Vol. 404. Austin, TX: Pro-ed, 1999.

[14] Gargiulo, R.M. (2006). *Special education in contemporary society: An introduction to*

exceptionality. Belmont, CA: Thomson Wadsworth

[15] Ho, Aileen K., et al. "Speech volume regulation in Parkinson's disease: Effects of implicit cues and explicit instructions." *Neuropsychologia* 37.13 (1999): 1453-1460.

[16] Mollaei, Fatemeh, Douglas M. Shiller, and Vincent L. Gracco. "Sensorimotor adaptation of speech in Parkinson's disease." *Movement Disorders* 28.12 (2013): 1668-1674.

[17] Kalinowski, Joseph, et al. "Effects of alterations in auditory feedback and speech rate on stuttering frequency." *Language and Speech* 36.1 (1993): 1-16.

[18] Hämäläinen, Matti, et al. "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain." *Reviews of modern Physics* 65.2 (1993): 413.

[19] Hall, Emma L., et al. "The relationship between MEG and fMRI." *Neuroimage* 102 (2014): 80-91.

[20] Singh, Sanjay P. "Magnetoencephalography: basic principles." *Annals of Indian Academy of Neurology* 17.Suppl 1 (2014): S107.

[21] Barkley, Gregory L., and Christoph Baumgartner. "MEG and EEG in epilepsy." *Journal of clinical neurophysiology* 20.3 (2003): 163-178.

[22] Knake, S., et al. "The value of multichannel MEG and EEG in the presurgical evaluation of 70 epilepsy patients." *Epilepsy research* 69.1 (2006): 80-86.

[23] Wipf, David P., et al. "Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG." *NeuroImage* 49.1 (2010): 641-655.

[24] Vrba, Jiri, and Stephen E. Robinson. "Signal processing in magnetoencephalography." *Methods* 25.2 (2001): 249-271.

[25] Wipf, David, and Srikantan Nagarajan. "A unified Bayesian framework for MEG/EEG source imaging." *NeuroImage* 44.3 (2009): 947-966.

[26] Wipf, David P., et al. "Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG." *NeuroImage* 49.1 (2010): 641-655.

[27] Owen, Julia P., et al. "Performance evaluation of the Champagne source reconstruction algorithm on simulated and real M/EEG data."*Neuroimage* 60.1 (2012): 305-323.

[28] Thompson, Megan. *Speech and Nonspeech Production I the Absence of the Vocal Tract.* Diss. University of California, San Francisco, 2018.

[29] Rabiner, Lawrence R., and Ronald W. Schafer. "Matlab exercises in support of teaching digital speech processing." *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014.

[30] Kewley-Port, Diane, and Charles S. Watson. "Formant-frequency discrimination for isolated English vowels." *The Journal of the Acoustical Society of America* 95.1 (1994): 485-496.

[31] Hillenbrand, James, et al. "Acoustic characteristics of American English vowels." *The Journal of the Acoustical society of America* 97.5 (1995): 3099-3111.

[32] Peterson, Gordon E., and Harold L. Barney. "Control methods used in a study of the vowels." *The Journal of the acoustical society of America* 24.2 (1952): 175-184.

[33] Sjölander, Kåre, et al. "Web-based educational tools for speech technology." *Fifth International Conference on Spoken Language Processing*. 1998.

[34] Sjölander, Kåre, and Jonas Beskow. "Wavesurfer-an open source speech tool." *Sixth International Conference on Spoken Language Processing*. 2000.

[35] Dalal, Sarang S., et al. "MEG/EEG source reconstruction, statistical evaluation, and visualization with NUTMEG." *Computational intelligence and neuroscience* 2011 (2011).

[36] Assaneo, M. Florencia, and David Poeppel. "The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm." *Science advances* 4.2 (2018): eaao3842.

[37] Zatorre, Robert J., Joyce L. Chen, and Virginia B. Penhune. "When the brain plays music: auditory–motor interactions in music perception and production." *Nature reviews neuroscience* 8.7 (2007): 547.

[38] Wernicke C. Der Aphasische Symptomencomplex. Breslau: Cohn and Weigert; 1874.

[39] Phillips DP, Farmer ME. Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. Behav Brain Res. 1990;40:85–94.

[40] Tallal P, Miller S, Fitch RH. Neurobiological basis of speech: a case for the preeminence of temporal processing. Ann N Y Acad Sci. 1993;682:27–47.

[41] Meyer M, Zaehle T, Gountouna VE, Barron A, Jancke L, Turk A. Spectro-temporal processing during speech perception involves left posterior auditory cortex. Neuroreport. 2005;16:1985–1989.

[42] Zaehle T, Wustenberg T, Meyer M, Jancke L. Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. Eur J Neurosci. 2004;20:2447–2456.

[43] Scott SK, Blank CC, Rosen S, Wise RJ. Identification of a pathway for intelligible speech in the left temporal lobe. Brain. 2000;123(Pt 12):2400–2406.

[44] Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. Neural substrates of phonemic perception. Cereb Cortex. 2005;15:1621–1631.

[45] Obleser J, Zimmermann J, Van Meter J, Rauschecker JP. Multiple stages of auditory speech perception reflected in event-related FMRI. Cereb Cortex. 2007;17:2251–2257.

[46] Rodd JM, Davis MH, Johnsrude IS. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb Cortex. 2005;15:1261–1269.

[47] Wagner AD, Pare-Blagoev EJ, Clark J, Poldrack RA. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron. 2001;31:329–338.

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.

Please sign the following statement:

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

Author Signature

08/15/2018

Date