12-2013

NEURAL CORRELATES OF AUDIOVISUAL SPEECH PERCEPTION IN APHASIA AND HEALTHY AGING

Sarah H. Baum

NEURAL CORRELATES OF AUDIOVISUAL SPEECH PERCEPTION IN APHASIA AND HEALTHY AGING

by

Sarah Haller Baum, B.S.

APPROVED:

_____
Michael Beauchamp, Ph.D.
Supervisory Professor

_____
Ruth Heidelberger, M.D., Ph.D.

_____
Harel Shouval, Ph.D.

_____
Michael Beierlein, Ph.D.

_____
Kartik Venkatachalam, Ph.D.

APPROVED:

_____
Dean, The University of Texas
Graduate School of Biomedical Sciences

NEURAL CORRELATES OF AUDIOVISUAL SPEECH PERCEPTION IN APHASIA AND HEALTHY AGING

A DISSERTATION

Presented to the Faculty of The University of Texas Health Science Center at Houston and The University of Texas M. D. Anderson Cancer Center Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Sarah Haller Baum, B.S., Houston, Texas

December, 2013

ACKNOWLEDGMENTS

My first thanks go to my advisor, Michael Beauchamp, for allowing me enough independence to encounter my own successes and failures while still giving me enough feedback to continually move forward in my research. To my labmates, thank you for all the helpful, supportive, and funny conversations, and for obliging me in an unknown number of coffee runs.

To the faculty of my advisory, examining, and supervisory committees, thank you for taking the time to provide thoughtful feedback and critiques of my experiments. Thank you to Ruth Heidelberger and Harel Shouval for serving on every single committee from the very beginning.

I would like to thank the Gulf Coast Consortia for the Theoretical and Computational Neuroscience Fellowship, which provided financial support and a number of resources for professional development.

The Neuroscience Program has been a great source of camaraderie and moral support throughout my time here. Many thanks especially to Heather Turner for sharing in my joy on the best days and providing support on the rough days, and to Audrey Nath, who continues to be a great friend and mentor.

Lastly, I would never be here without my parents, who knew what I was capable of before I did and never let me settle for anything less. And to my siblings, Aron, Emma, and Hannah, who despite being younger than me are the most fiercely protective and supportive people I know. Thank you for letting me spend time halfway across the country to go follow my dream.

NEURAL CORRELATES OF AUDIOVISUAL SPEECH PERCEPTION IN APHASIA AND HEALTHY AGING

Sarah Haller Baum, B.S.

Supervisory Professor: Michael Beauchamp, Ph.D.

Understanding speech in face-to-face conversation utilizes the integration of multiple pieces of information, most importantly the auditory vocal sounds and visual lip movements. Prior studies of the neural underpinnings of audiovisual integration in the brain have provided converging evidence to suggest that neurons within the left superior temporal sulcus (STS) provide a critical neural hub for the integration of auditory and visual information in speech. While most studies of audiovisual processing focus on neural mechanisms within healthy, young adults, we currently know very little about how changes to the brain can affect audiovisual integration in speech. To examine this further, two particular cases of changing neural structure were investigated. I first conducted a case study with patient SJ, who suffered damage from a stroke that injured a large portion of her left tempo-parietal area, including the left STS. I tested SJ five years after her stroke with behavioral testing and determined that she is able to integrate auditory and visual information in speech. In order to understand the neural basis of SJ's intact multisensory integration abilities, I examined her and 23 age-matched controls with functional magnetic resonance imaging (fMRI). SJ had a greater volume of multisensory cortex as well as greater response amplitude in her right STS in response to an audiovisual speech illusion than the

age-matched controls. This evidence suggests that SJ's brain reorganized after her stroke such that the right STS now supports the functions of the stroke damaged left-sided cortex. Because changes to the brain occur even with healthy aging, I next examined the neural response to audiovisual speech in healthy older adults. Many behavioral studies have noted that older adults show not only performance declines during various sensory and cognitive tasks, but also greater variability in performance. I sought to determine if there is a neural counterpart to this increased behavioral variability. I found that older adults exhibited greater intrasubject variability in their neural responses across trials compared to younger adults. This was true in individual regions-of-interest in the multisensory speech perception network and across all brain voxels that responded to speech stimuli. This increase in variability may underlie a decreased ability of the brain to distinguish between similar stimuli (such as the categorical boundaries of speech perception), which could link these findings to declines in speech perception in aging.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

ABSTRACT iv

LIST OF FIGURES vii

LIST OF TABLES viii

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: MULTISENSORY SPEECH PERCEPTION WITHOUT THE LEFT SUPERIOR TEMPORAL SULCUS 11

- Introduction 12
- Methods 13
- Results 20
- Conclusions 33

CHAPTER 3: INCREASED NEURAL VARIABILITY IN MULTISENSORY SPEECH PERCEPTION IN OLDER ADULTS 36

- Introduction 37
- Methods 39
- Results 50
- Conclusions 64

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS 70

BIBLIOGRAPHY 77

VITA 94

LIST OF FIGURES

Figure 1.1 Types of speech stimuli. 3

Figure 1.2 Structural MRI of a healthy young adult subject. 4

Figure 2.1 Anatomical MRI of SJ 14

Figure 2.2 Stimuli used in experiments. 15

Figure 2.3 Behavioral testing results 23

Figure 2.4 fMRI activation during localizer scan 25

Figure 2.5 Multisensory responses in the STS in SJ and controls 28

Figure 2.6 Hemodynamic response to all audiovisual stimuli 32

Figure 3.1 BOLD responses to audiovisual speech 51

Figure 3.2 Intrasubject variability in older and younger subjects 53

Figure 3.3 Whole-brain analysis of differences in older and younger adults 59

Figure 3.4 Relationship between head motion and variability 62

LIST OF TABLES

Table 2.1 Anatomical regions in left hemisphere impacted by stroke. 21

Table 2.2 Areas of differential activation in SJ and controls 30

Table 3.1 Activation to audiovisual speech syllables across both younger and older adults. 54

Table 3.2 Whole-brain response amplitude in younger and older adults. 55

Table 3.3 Whole-brain intersubject variability in younger and older adults. 56

Table 3.4 Whole-brain intrasubject variability in younger and older adults. 57

CHAPTER 1: INTRODUCTION

Speech perception is one of the most important cognitive functions performed by the human brain. In face-to-face conversation, understanding speech is a multisensory process in which auditory information (vocal sounds) and visual information (lip movements) are integrated into a single, coherent percept. Although these two pieces of information are naturally and automatically integrated in the brain when both cues are clear and salient, the combination of both pieces of information is even more important when either cue is presented in the context of noise, such as in a loud room (Sumby & Pollack, 1954, MacLeod & Summerfield, 1990, Ross et al., 2007).

To measure multisensory integration in a laboratory setting, multiple types of speech stimuli are used. Speech can be presented with auditory information alone, which in most situations is clearly and accurately understood (Figure 1.1A). Typical, everyday speech includes congruent auditory and visual information (Figure 1.1B). One obstacle to studying multisensory integration in speech, however, is that with clear, natural speech, there is often no difference in the reported percept between auditory-only and audiovisual presentation of speech (compare percepts in Figure 1.1A and B). Therefore, although a listener will integrate typical congruent audiovisual speech if they pay attention to both cues, it is harder to show on a trial-to-trial basis that subjects were successfully integrating both auditory and visual information because the same percept could be achieved by only listening to the auditory component of the stimulus.

Consequently, a third type of speech stimulus is used in many audiovisual speech experiments. The McGurk effect is an audiovisual speech illusion

(McGurk & MacDonald, 1976) in which a video of a voiced syllable (*e.g.* “ba”) is paired with a different mouthed syllable (*e.g.* “ga”), which for some subjects results in the perception of a completely different, third syllable (*e.g.* “da”) (Figure 1.1C). The only way to account for this perception is through multisensory integration, since the percept is different than the actual presented sensory information in either the auditory or visual component. Using this incongruent, non-natural stimulus provides a clear marker for audiovisual integration in speech on an individual trial basis, and is therefore a powerful tool for studying multisensory integration.

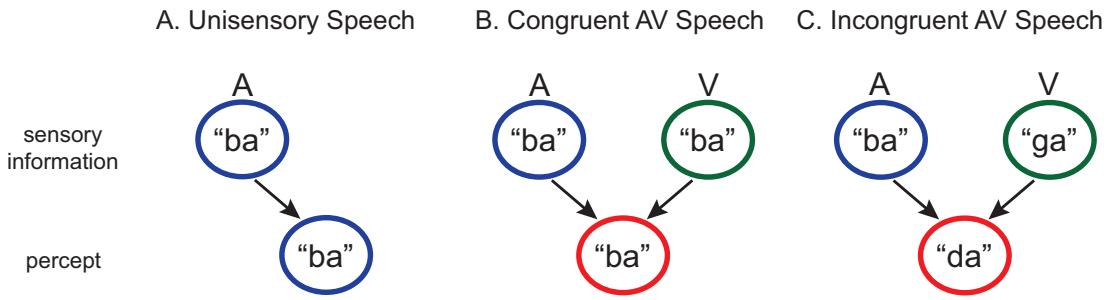

Figure 1.1 Types of speech stimuli.
A: Unisensory speech contains only auditory information presented without any visual information. The resulting percept is driven by the auditory information alone. **B:** Congruent AV speech contains both auditory and visual information with the same syllable presented in each component. The resulting percept integrates both pieces of information. **C:** Incongruent audiovisual speech contains both auditory and visual information but the auditory syllable and visual syllable are different. In some combinations, the two syllables are integrated into a fused percept.

The general principles of multisensory integration in speech have been studied extensively through behavioral tests alone, but the combination of behavioral data and neuroimaging can shed light on the neural mechanisms that support multisensory integration. The most popular technique to measure human brain function non-invasively is functional magnetic resonance imaging (fMRI) (Friston, 2009). Unlike other brain imaging techniques, MRI can be used to image both brain structure *and* function. These two complementary pieces of evidence provide information about both neural structure and neural activity, both of which may differ between groups of subjects and patients.

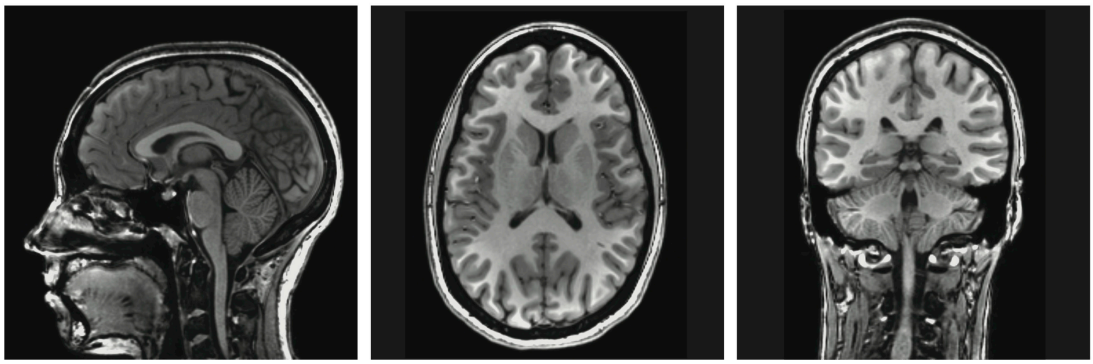

Figure 1.2 Structural MRI of a healthy young adult subject.
Sagittal, axial, and coronal slices from a structural T_1 scan of a healthy subject. Images taken from MRI scan of author.

Structural MRI images the underlying neural anatomy of the brain. In the presence of a strong magnetic field (B_0), protons in hydrogen atoms align with the direction of the magnetic field. A radio frequency pulse then temporarily knocks these protons out of alignment. The most common structural scan, T_1, measures the time it takes these protons to relax back to the lower energy state and realign with the magnetic field. This time constant is different for various

biological tissues (*e.g.*, white matter, gray matter, bone, and cerebrospinal fluid). For example, protons in gray matter realign much slower than white matter, which results in the different level of brightness between the two tissue types: in a T_1 scan, gray matter is darker than white matter (Huettel et al., 2009) (Figure 1.2). The frequency at which protons are excited depends on the strength of the magnetic field. Therefore, by systematically varying the exact strength of the magnetic field in space along x, y, and z gradients, a single slice of brain can be excited at a time, which allows for the measurement of relaxation rates (and therefore different tissue types) in 3D space. A typical T_1 scan lasts approximately 4 minutes and has a resolution of 1mm x 1mm x 1mm.

In contrast, functional MRI (fMRI) relies on the local changes in blood flow induced by increased neural activity in a given region. Following increased neuronal activity, neurons send signals locally to nearby blood vessels to increase the blood flow and bring more oxygenated blood. These signals increase blood flow and overcompensate with an excess of oxygenated blood (Fox & Raichle, 1986, Fox et al., 1988). This response is called functional hyperemia, although the exact physiological mechanisms of this response are unknown. Astrocytes have been strongly implicated in the link between neuronal activity and the resulting increases in cerebral blood flow (Gordon et al., 2007), but recent work suggests that stimulus-induced vasodilation can occur even without calcium-dependent release of vasodilators (Nizar et al., 2013). Oxygenated blood is diamagnetic, meaning that it does not create any magnetic moment because there are no unpaired electrons. Conversely, deoxygenated

blood is paramagnetic, which creates small distortions in the magnetic field because it has a significant magnetic moment and unpaired electrons. Because oxygenated blood and deoxygenated blood have different magnetic susceptibilities, the proportion of oxygenated blood and deoxygenated blood can be detected in a high strength magnetic field (Ogawa & Lee, 1990). The contrast of oxygenated and deoxygenated blood is known as the blood oxygen level dependent (BOLD) signal and is useful to study neural processing because the proportion of oxygenated blood changes with changes in brain activity. A second time constant used in MRI is the T_2 constant. Unlike the T_1 constant, which measures how long protons take to realign to the longitudinal B_0 plane, T_2 measures relaxation in the transverse magnetic plane (perpendicular to the B_0 field). When inhomogeneities are present (such as those disturbances created by paramagnetic deoxygenated blood), the decay constant is known as T_2^*. A T_2^* scan is one of the most commonly used scans for measuring functional brain activation in MRI (Logothetis & Wandell, 2004). In a typical T_2* scan, an entire brain volume is acquired approximately every 2 seconds. The resolution of an fMRI scan varies depending on the strength of the magnetic field. With a 3T scanner at the UT Medical School a resolution of 2.75mm x 2.75mm x 2.75mm can be obtained.

Despite the many advantages that MRI offers, there are also some limitations. The BOLD signal is both an indirect and relative measure of neuronal activity. Neural activity is measured in relative units, percent signal change from baseline, and a typical significant difference in neural activity will be in the range

of 0.1% - 1% signal change. Unlike neuronal activity, which changes on a millisecond time scale, the hemodynamic signal is much slower and peaks approximately 4-6 seconds after the presentation of a stimulus. Due to the slow nature of the BOLD signal, fMRI is not well suited to research questions involving the exact timing of different neural inputs and processes. Because of the relatively low spatial resolution (~3mm), activity is averaged over thousands of cells. Studies of the physiological basis of fMRI have used simultaneous measurements of electrophysiological and fMRI responses from monkeys. The results of these studies suggest that the BOLD response amplitude represented a combination of local field potentials and action potentials (Logothetis & Wandell, 2004). Roughly 90% of synapses in cortex are excitatory, so the majority of this average comes from excitatory signals (Braitenberg & Schuez, 1998).

By combining fMRI and behavioral measures, we can investigate the neural substrates of behavior and sensory processing. Previous studies have suggested that the superior temporal sulcus (STS) is a critical site for the integration of audiovisual speech (Calvert et al., 2000, Miller & D'Esposito, 2005). The STS exhibits a greater BOLD signal response to multisensory stimuli as compared to auditory-only or visual-only stimuli (Beauchamp et al., 2004, Stevenson & James, 2009). The amplitude of the left STS is correlated with the amount audiovisual integration in individual subjects, as measured by susceptibility to the McGurk effect, in both children and young adults (Nath et al.,

2011, Nath & Beauchamp, 2012). Together these findings suggest that the STS is a critical locus for the integration auditory and visual information.

In the experiments presented in this dissertation, a combination of behavioral and neuroimaging measures are used to examine neural correlates of multisensory speech perception. The neural structures that support the function of multisensory integration in the brain are well studied in healthy, young subjects. However, what happens if these underlying neural structures change? I have conducted two studies that examine pieces of this puzzle. The goals of this project were two-fold:

1. To describe a case study of audiovisual processing in which a stroke patient incurred damage to the left STS (Chapter 2).
2. To characterize changes to neural responses to multisensory speech perception in healthy older subjects (Chapter 3).

I report a patient, SJ, who suffered a cerebral vascular accident that damaged the left tempo-parietal area, including the left STS, resulting in mild anomic aphasia (Baum et al., 2012). Previous studies have demonstrated a critical role of the left STS in multisensory speech perception (Scott & Johnsrude, 2003, Beauchamp, 2005, Miller & D'Esposito, 2005, Stevenson & James, 2009, Nath & Beauchamp, 2011, 2012). Because temporary disruption of the left STS with TMS impairs multisensory speech perception (Beauchamp et al., 2010), one might expect the lesion suffered by SJ to greatly reduce multisensory integration. I first tested patient SJ with a series of behavioral tests to determine if she was able to integrate auditory and visual information in

speech. I predicted that, given the damage to the left STS, SJ's multisensory abilities should be at the same level or worse than the controls whose left STS activity was temporarily disrupted with TMS. Surprisingly, SJ demonstrated intact multisensory abilities in the behavioral testing.

Based on the observed improvements in speech perception, neural plasticity and rehabilitation in SJ might have resulted changes in neural processing, leading to her improved abilities. This would predict different patterns of brain activity during multisensory speech perception in SJ compared with age-matched controls. To test this hypothesis, I then completed an fMRI experiment with SJ and 23 healthy age-matched controls to determine the neural correlates of her multisensory abilities. I found that SJ had more functionally multisensory cortex in her right STS than any of the age-matched controls tested. Furthermore, the response amplitude in the right STS to McGurk stimuli was much greater in SJ than in healthy controls. Together, these results suggest that the SJ's right STS now suberves the multisensory functions previously completed by the left STS.

I next considered the healthy older adults as a cohort to investigate the neural basis of changes in audiovisual perception during the process of aging. Older adults experience declines in their unisensory abilities, notably their ability to hear speech in noise, even without any significant hearing loss (Divenyi et al., 2005). Not only does performance decrease with aging, there is also a significant amount of performance variability, such that performance is not only worse but also more inconsistent from trial to trial. I hypothesized that the well-

documented increases in behavioral variability could also be observed on a neural level. To test this, I presented multiple trials of simple audiovisual speech stimuli and measured each subject's variability from their individual neural responses. I found that older adults have much more inconsistent neural responses to speech (higher intrasubject variability) from trial-to-trial than younger adults.

Increased neural variability may provide a mechanism for declines in speech perception observed in healthy aging. This increased variability may lead to difficulty in discriminating between stimuli, which would decrease the ability to accurately identify sensory information. Many studies analyze only the differences in response amplitudes, but these results demonstrate that it may be important to examine variability (as well as average amplitude of the response) when examining changes in neuronal processing in the aging brain. These patterns of differential variability as a function of age or cognitive ability may provide an additional way to characterize changes in brain function that are often overlooked.

CHAPTER 2: MULTISENSORY SPEECH PERCEPTION WITHOUT THE LEFT SUPERIOR TEMPORAL SULCUS

Introduction

Speech can be understood through the auditory modality alone, but combining audition with vision improves speech perception (Sumby & Pollack, 1954, Stein & Meredith, 1993, Grant & Seitz, 2000). One striking behavioral example of audiovisual multisensory integration in speech perception is the McGurk effect (McGurk & MacDonald, 1976) in which an auditory syllable paired with a video clip of a different visual syllable results in the percept of a distinct new syllable (*e.g.* auditory “ba” + visual “ga” results in the percept “da”). Because the fused percept is different than either the auditory or visual stimulus, it can only be explained by multisensory integration.

A number of studies suggest that the left superior temporal sulcus (STS) is an important site of audiovisual multisensory integration. The left STS exhibits a larger BOLD response to multisensory stimuli as compared to unisensory stimuli (Calvert et al., 2000, Beauchamp et al., 2004, Stevenson & James, 2009). Tracer studies in rhesus macaque monkeys reveal that the STS is anatomically connected both to auditory cortex and extrastiate visual cortex (Seltzer et al., 1996, Lewis & Van Essen, 2000). There is a correlation between the amplitude of activity in the left STS and the amount of McGurk perception in both individual adults (Nath & Beauchamp, 2012) and children (Nath et al., 2011). Inter-individual differences in left STS activity have also been linked to language comprehension abilities (McGettigan et al., 2012). When the left STS is temporarily inactivated with transcranial magnetic stimulation (TMS) in normal subjects, the McGurk effect is reduced (Beauchamp et al., 2010). Unlike the transient disruptions created by TMS, lesions caused by brain injury can give insight into the results of brain plasticity that occur after a stroke. In particular, damage to areas in the language network can result in brain reorganization, with increased activity in the areas homologous to the damaged tissue (Buckner et al., 1996, Thomas, 1997, Cao et al., 1999, Blasi et al., 2002, Winhuisen et al., 2005).

We describe a patient, SJ, with a lesion that completely ablated her left posterior STS. Following her stroke, SJ underwent intensive behavioral therapy. In the years following her stroke, her speech perception abilities improved. Five years after her stroke SJ demonstrated multisensory speech perception similar to 23 age-matched controls when tested with two independent behavioral measures. To understand the neural substrates of this ability, we examined patient SJ and age-matched controls with structural and functional MRI.

Methods

Patient SJ

All subjects provided informed consent under an experimental protocol approved by the Committee for the Protection of Human Subjects of the University of Texas Health Science Center at Houston. All participants received compensation for their time. Patient SJ is a 63 year-old, right-handed female who presented with a language impairment following a stroke, which destroyed a large portion of her left temporal lobe, including the left STS (Figure 2.1 and Table 2.1). Patient SJ was 58 years old when she suffered a stroke in the left tempo-parietal area in September 2006. Prior to her stroke SJ worked in public relations and had completed one year of college. SJ's performance on the Western Aphasia Battery indicated a classification of anomic aphasia. Her auditory comprehension was impaired 3 years after the stroke (48% on auditory lexical decision and 86% for CV miminal pairs, compared with expected 95 – 100% for controls). 5 years after the stroke, her auditory recognition had improved to near normal range (87% on auditory lexical decision and 95% for CV miminal pairs). SJ was scanned two times, once for structural MRI in February 2010, and again for structural and functional MRI in March 2011.

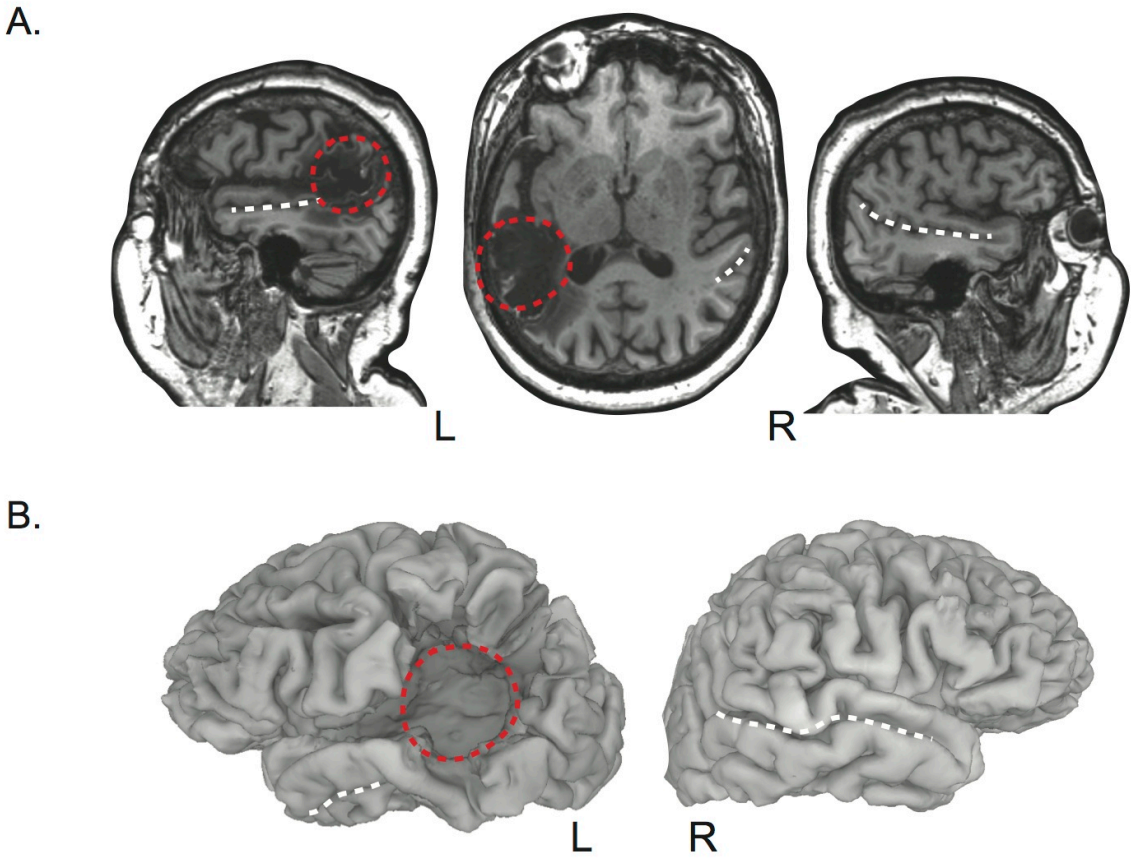

Figure 2.1 Anatomical MRI of SJ
A. Sagittal and axial slices of SJ's structural MRI. White dashed lines indicate the location of the STS. Red dashed circle indicates stroke-damaged cortex in left hemisphere (left is left on all images).
B. Cortical surface reconstruction of SJ's brain from the structural MRI.
Taken from Baum et al. (2012).

Healthy Age-Matched Control Subjects

23 healthy older adults ranging in age from 53-75 years (14 female, mean age 62.9 years) served to provide a healthy age-matched comparison to patient SJ. Participants were recruited through word-of-mouth and flyers distributed around the greater Houston area. 21 subjects were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). All subjects were fluent English speakers.

Stimuli used for testing

Stimuli consisted of a digital video recording of a female native English speaker speaking “ba”, “ga”, “da”, “pa”, “ka” and “ta”. Digital video editing software (iMovie, Apple Computer) was used to crop the total length of each video stimulus such that each clip both started and ended in a neutal, mouth-closed position. Each video clip ranged from 1.7 to 1.8 seconds.

Auditory-only stimuli were created by extracting the auditory track of each video and pairing it with white visual fixation crosshairs on a gray screen. Visual-only stimuli were created by removing the auditory track of each video. Two separate McGurk stimuli were created by pairing the auditory “ba” with the visual of “ga” (canonical percept “da”), and pairing the auditory “pa” with the visual of “ka” (canonical percept “ta”). Non-McGurk incongruent stimuli were created by reversing the pairing of the two McGurk stimuli (auditory “ga” with visual “ba”, resulting in the percept “ga”, and auditory “ka” with visual “pa”, resulting in the percept “ka”). These stimuli were used for both behavioral testing and the fMRI experiment (Figure 2.2). Eight additional McGurk stimuli were obtained from youtube.com for additional behavioral testing with SJ.

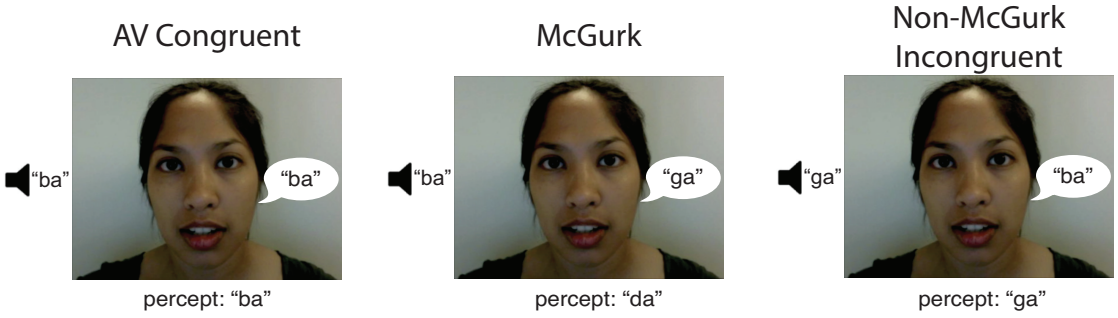

Figure 2.2 Stimuli used in experiments.
Three audiovisual (AV) speech stimuli used in the experiments described in Chapters 2 and 3. **A:** AV Congruent stimuli: same voiced and mouthed syllables. Incongruent stimuli (spoken and mouthed syllables different) were categorized as either **B:** McGurk stimuli (likely to be integrated and result in a fused percept) or **C:** Non-McGurk Incongruent stimuli (not integrated, percept likely to be auditory syllable).

Behavioral Testing of Healthy Controls

Each subject's perception of auditory only, congruent, and McGurk syllables was assessed. Stimuli were presented in two separate runs: auditory-only syllables (10 trials of each syllable) and AV syllables (10 trials each of "ba"/"da" McGurk syllables, "pa"/"ka" McGurk syllables, and "ba", "da", "pa" and "ka" congruent syllables) in random order. Auditory stimuli were delivered through headphones at approximately 70 dB, and visual stimuli were presented on a computer screen. For all stimuli, subjects were instructed to watch the mouth movements (if present) and listen to the speaker. Perception was assessed by asking the subject to verbally repeat out loud the perceived syllable. The response was open choice and no constraints were placed on allowed responses. This response format was chosen because it has been shown to provide a more conservative estimate of McGurk perception (Colin et al., 2005). All spoken responses were recorded by a microphone and the experimenter writing down each response. For SJ, the testing procedure was identical, but additionals trials of McGurk stimuli were presented (15 trials *vs.* 10 in controls).

Morphed Audiovisual Syllables

An additional, independent, test of multisensory integration was obtained by measuring SJ's perception of audiovisual syllables along a continuum of "ba" to "da" (Massaro et al., 1993). Synthetic auditory speech stimuli were created by taking tokens of "ba" and "da" and manipulating the first 80ms to create five auditory syllables ranging from A1 (100% ba/0% "da") to A5 (0% "ba"/100% "da"). Similarly, synthetic visible speech stimuli were created by using a computer-animated display whose mouth position at the syllable onset was systematically altered to create V1 (100% "ba"/0% "da") to V5 (0% "ba"/100% "da"). Each audiovisual syllable stimulus (five auditory times five visual for 25 total) was presented 20 times in random order in a two alternative forced choice task where SJ was instructed to respond if she perceived the audiovisual syllable to be more like "ba" or "da". Responses were made on a mouse with the left button labeled "ba" and the right button labeled "da". Written instructions were also presented on the screen after each trial. We compared SJ's responses with those of 82 healthy subjects viewing the same stimuli, reported in Massaro et al. (1993).

fMRI McGurk Experiment

Each fMRI run lasted approximately four minutes and two scan runs were collected from each subject. In each run, single syllables were presented within the 2-second trial using a rapid event-related design. Trials were pseudo-randominzed for an optimal rapid-event related order (Dale, 1999). In each trial, a video clip was presented followed by fixation crosshairs for the remainder of the trial. The crosshairs were positioned such that they were in the same location as the mouth during visual speech in order to minimize eye movements and draw attention to the visual mouth movements. Subjects responded to target trials only (the word "press"). For SJ and six

control subjects, each run contained 25 McGurk trials, 25 non-McGurk incongruent trials, 25 congruent trials, 20 target trials, and 25 trials of fixation baseline. For the remaining 17 control subjects each run contained 40 McGurk trials, 20 non-McGurk incongruent trials, 20 congruent trials, 15 target trials and 25 trials of fixation baseline. All stimuli were identical to those used for behavioral testing outside the scanner.

fMRI Functional Localizer Experiment

In order to prevent bias when analyzing the McGurk fMRI data, a separate scan series was performed to identify independent regions of interest. The functional localizer scan consisted of six blocks of one syllable words (two auditory-only, two visual-only and two audiovisual blocks in random order) which contained 20 seconds of stimulus (10 two second trials, one word per trial) followed by 10 seconds of fixation baseline between each block. Each block contained a target trial (the word “press”) of the same stimulus type as the other stimuli in the block; subjects were instructed to pay attention to each stimulus and press a button during target trials but not to any other stimuli.

MRI and fMRI Analysis

Two T1-weighted MP-RAGE anatomical MRI scans were collected at the beginning of each scanning session with a 3 tesla whole-body MR scanner (Phillips Medical Systems) using a 32-channel head coil. The two anatomical scans were aligned to each other and averaged in order to provide maximal gray-white matter contrast. These scans were then used to create a cortical surface model using FreeSurfer (Dale et al., 1999, Fischl et al., 1999) for visualization in SUMA (Argall et al., 2006). For the fMRI scan series, T2* weighed images were collected using gradient echo-planar imaging (TR = 2000 ms, TE = 30 ms, flip angle = 90°) with in-plane resolution of 2.75 x 2.75 mm. The McGurk syllable scan series and localizer scan series consisted of 123 and 138 brain volumes, respectively. The first three volumes were discarded because they were collected before equilibrium magnetization was reached. This resulted in 120 and 135 usable brain volumes, respectively. Auditory stimuli were presented through MRI-compatible in-ear headphones (Sensimetrics, Malden, MA) which were covered with ear muffs to reduce the amount of noise from the scanner. Visual stimuli were presented on a projection screen with an LCD projector and viewed through a mirror attached to the head coil. Responses to the target trials were collected using a fiber-optic button response pad (Current Designs, Haverford, PA).

Analysis of the functional scan series was conducted using Analysis of Functional NeuroImages (AFNI) (Cox, 1996). Data were analyzed for each subject individually and then the data for all healthy control subjects was combined using a random-effects model. Functional data for each subject was first aligned to the averaged anatomical dataset and then motion-corrected using a local Pearson correlation (Saad et al., 2009). The analysis

of all voxels was carried out with the AFNI function *3dDeconvolve,* which uses a generalized linear model utilizing a maximum-likelihood approach. Tent-zero functions were used in the deconvolution to estimate the individual hemodynamic response function in each voxel for each stimulus type that began at stimulus onset and ended 16 seconds after stimulus onset for rapid event related runs and 26 seconds for block design runs.

A modified, conservative t-test (Crawford, 1998) was used to compare single data points from SJ with averaged data from controls. To test for the significance of any differences in fMRI response amplitude by stimulus type, the within type variance was computed as follows. For controls, we considered the average response to a stimulus in each individual control subject as a sample. For SJ, we considered the response to individual presentations of each stimulus, calculated with a least-square sum method in the AFNI program *3dLSS* (Mumford et al., 2012). This analysis was used for all ROIs except for the left STS, for which the response was 0 for all trials, necessitating the use of the conservative single point t-test.

Group Analysis

Two strategies were used for group analysis. Converging evidence from both strategies indicates a robust difference between SJ and controls. In the first strategy, regions of interest (ROI) are selected based on the individual anatomy in each subject (Saxe et al., 2006). Because the course of the STS is highly variable across subjects, standard 3-D anatomical templates fail to accurately align STS gray matter. Using a cortical-surface based analysis, the STS in each subject is aligned to the STS of a 2-D template for labeling purposes. This allows for unbiased measurement of activity in the STS (and other regions). Each ROI was created using the FreeSurfer anatomic parcellation of the cortical surface constructed from each individual subject's structural scans (Fischl et al., 2004, Destrieux et al., 2010). The parcellation defined 74 distinct regions for each hemisphere in each subject. SJ's automated parcellation was manually inspected to ensure that the 3-D reconstruction was an accurate representation of her structural damage. This parcellation was then manually edited for SJ's left hemisphere to ensure that no labels were assigned to the lesion zone.

ROIs created in each subject's individual native space were used in the main analysis, thus any potential discrepancy between the un-normalized brain and reference template did not affect the analysis results. These ROIs were then analyzed with data from independently collected runs, eliminating bias (Kriegeskorte et al., 2009). The STS ROI was defined by finding all voxels in the posterior half of the anatomically defined STS that responded to both auditory-only words and visual-only words ($t > 2$ for each modality). For some subjects ($n = 5$ in left hemisphere, $n = 1$ in right hemisphere), there were no voxels in the posterior STS that were significantly active during both auditory-only and visual-only word blocks. For these subjects the STS ROI was defined by finding all voxels in the anatomically defined posterior STS

that were active (t > 2) during the audiovisual word blocks. The auditory cortex ROI was defined by finding voxels in the anatomically parcellated transverse temporal gyrus, lateral superior temporal gyrus and planum temporale that were active (t > 2) during the auditory-only blocks. The extrastriate visual cortex ROI was defined by finding voxels in the anatomically parcellated extrastriate lateral occipitotemporal cortex that were active (t > 2) during the visual-only blocks. We chose a later visual area to study because of its prominent role in visual speech perception and strong activation during audiovisual speech.

In the second strategy, a whole-brain voxel-wise analysis is used (Friston et al., 2006). Each individual subject brain and functional dataset was aligned to the N27 atlas brain (Mazziotta et al., 2001) with the *auto_tlrc* function in AFNI. The functional dataset for each subject was then smoothed using a 3 x 3 x 3mm FWHM Gaussian kernel. We wished to minimize blurring between the ROIs of interest and adjacent ROIs, so a small blurring kernel of approximately the same size as the voxel was chosen (Skudlarski et al., 1999). Areas with significantly different activation to McGurk stimuli between SJ and controls were searched for with *3dttest++.* These results were then transformed from the MRI volume to the cortical surface using *3dSurf2Vol* and clusters were identified with *SurfClust*. Clusters size threshold was 500 mm^2 with a z-score threshold of 3.5.

Results

Location and quantification of the lesion

Patient SJ's lesion destroyed a substantial portion of the lateral posterior left hemisphere (Figure 2.1 and Table 2.1). To quantify the extent of the lesion, we used automated anatomical parcellation to compare SJ's left hemisphere with 23 age-matched controls. The supramarginal gyrus and the STS were the areas with the greatest loss of gray matter. The lesion also extended into the temporal plane of the superior temporal gyrus, the location of auditory cortex.

Label	t-value	Delta (mm^3)	Volume in SJ (mm^3)	Mean ± SD Volume in Controls (mm^3)
Supramarginal gyrus	6.8	-4984	438	5422 ± 714
Superior Temporal Sulcus	5.5	-4872	3038	7910 ± 867
Postcentral sulcus	4.1	-2023	1376	3399 ± 482
Inferior segment of the circular sulcus of the insula	4.7	-1853	547	2400 ± 385
Temporal plane of the superior temporal gyrus	4.4	-1732	4	1736 ± 389
Posterior segment of the lateral fissure	6.6	-1376	14	1390 ± 203
Anterior transverse temporal gyrus	4.8	-856	30	886 ± 174
Long insular gyrus and central sulcus of the insula	4.5	-747	287	1034 ± 163
Transverse temporal sulcus	4	-429	4	433 ± 287

Table 2.1 Anatomical regions in left hemisphere impacted by stroke. Column 1 shows the FreeSurfer automatic parcellation anatomical label. Column 2 shows the t-value of the volume difference between SJ and controls. All differences are statistically significant at a level of $p < 0.01$ corrected for multiple comparisons. Column 3 shows the difference between the gray matter volume in SJ and the average gray matter volume in 23 age-matched controls (column 4 – column 5).
Taken from Baum et al. (2012).

Auditory and McGurk Perception: Behavioral Results

Sensory input is a prerequisite for multisensory integration. Because the lesion damaged regions of auditory cortex, we first examined SJ's auditory comprehension. When compared with 23 age-matched controls during our auditory-only syllable identification task, SJ was within the normal range (78% in SJ vs. 90% ± 15% in controls, $t_{22} = 0.75$, $p = 0.46$; Figure 2.3A). Next, we examined SJ's perception of McGurk stimuli, incongruent auditory and visual syllables in which an illusory percept indicates the presence of multisensory integration. SJ and controls reported similar rates of the illusory McGurk percept (66% vs. 59% ± 42%, $t_{22} = 0.16$, $p = 0.87$; Figure 2.3B).

Morphed Audiovisual Stimuli: Behavioral Results

As an independent test of multisensory integration, we presented 25 morphed audiovisual syllables along a continuum from "ba" to "da". SJ's perception was significantly influenced by both auditory and visual information. For instance, an ambiguous auditory stimulus (A4) was perceived as "da" 10% of the time when paired with one visual stimulus (V1) but was perceived as "da" 75% of the time when paired with a different visual stimulus (V5) ($p = 10^{-8}$ with binomial distribution). Conversely, an ambiguous visual stimulus (V4) was perceived as "da" 35% when paired with one auditory stimulus (A1) but 75% when paired with a different auditory stimulus (A5) ($p = 10^{-5}$ with binomial distribution). While SJ's multisensory integration in this task was significant, it was weaker for some stimuli than in the 82 controls tested by Massaro (1998) (A4V1, 10% vs. 66% ± 30% "da", $t_{81} = 1.91$, $p = 0.06$; A4V5, 75% vs. 98% ± 2%, $t_{81} = 9.38$, $p = 10^{-14}$; A1V4, 35% vs. 17% ± 25%, $t_{81} = 0.69$, $p = 0.49$; A5V4, 75% vs. 98% ± 2%, $t_{81} = 8.62$, $p = 10^{-13}$) (Figure 2.3C).

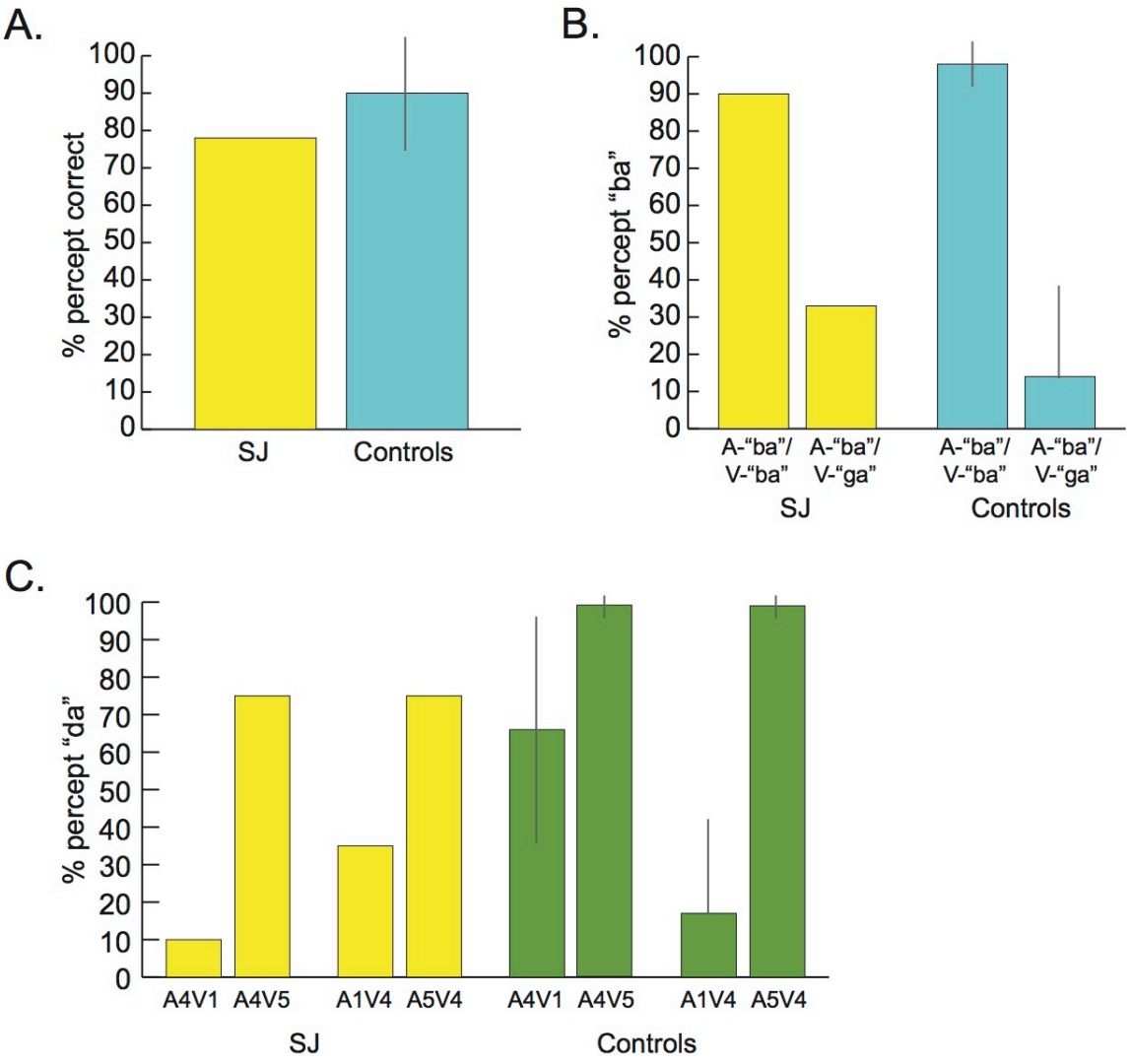

Figure 2.3 Behavioral testing results
A. Averaged auditory-only performance for six syllables (chance performance 17%) for SJ (yellow) and age-matched controls (blue).
B. Behavioral performance for one congruent audiovisual stimulus and one McGurk stimulus for SJ (yellow) and age-matched controls (blue).
C. Behavioral performance with 4 exemplar audiovisual morphed syllables. Data for SJ (yellow) and controls (green); control data from (Massaro, 1998).
Taken from Baum et al. (2012).

Functional MRI of Patient SJ and controls

SJ's behavioral results showed evidence for multisensory integration despite the extensive damage to her left STS. To understand the neural substrates of this preserved integration, we used fMRI to examine brain responses to multisensory speech.

We first presented separate blocks of auditory, visual and audiovisual words. Normal controls showed bilateral responses to audiovisual speech stimuli, with especially strong responses in the left superior temporal gyrus (STG) and STS. As expected from the extensive lesional damage, no activity was observed in SJ's left STS. However, activity was observed in her right hemisphere. Especially for the right STS, this activity appeared more extensive than in normal controls (Figure 2.4A). We used three strategies to quantify this observation. First, we measured the volume of active cortex within ROIs as defined by the localizer scan consisting of whole words. Second, we measured the amplitude of the response within localizer-defined ROIs to McGurk stimuli. Third, we performed a whole-brain analysis of activity evoked by the McGurk stimuli.

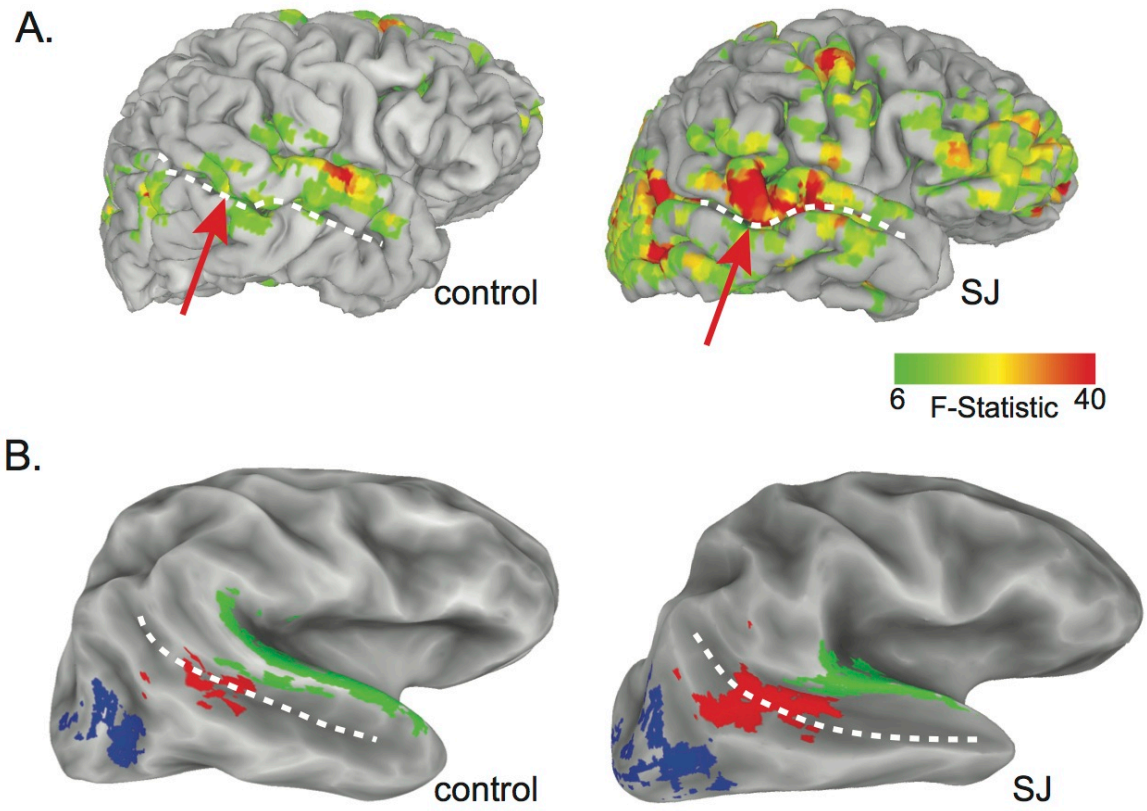

Figure 2.4 fMRI activation during localizer scan
A. Response to audiovisual speech in right hemisphere (lateral view of cortical surface model, color scale indicates significance of response) in one age-matched control (left, case IN) and stroke patient (right, case SJ). White dashed lines indicate STS, red arrow indicates activity in right STS.
B. Location of STS (red), extrastriate visual cortex (blue), and auditory cortex (green) ROIs in the right hemisphere of age-matched control (left, case IN) and stroke patient (right, case SJ).
Taken from Baum et al. (2012).

Method 1:Volume of Activated Cortex

To quantify activity, we measured the volume of cortex that showed significant responses to whole word audiovisual speech in three regions of interest: the STS, lateral extrastriate visual cortex, and auditory cortex (Figure 2.4B). As expected from the damage caused by the lesion, there was no active cortex in SJ's left STS vs. a large volume of STS activation in controls (0 vs. 34 ± 27 mm^3, t_{22}= 6.18, p = 10^{-6}) (Figure 2.5A). However, in right STS, SJ had much *more* active cortex than normal controls (96 vs. 30 ± 20 mm^3, t_{22} = 3.21 , p = 0.004). In fact, the volume of active cortex in SJ's right STS was greater than in any normal individual (Figure 2.5B). This finding (less active cortex in left hemisphere, more active cortex in right hemisphere) was not found in other ROIs. In extrastriate visual cortex, located close to the STS but just posterior and ventral to the lesion zone, there was no significant difference between SJ and controls in either the left hemisphere (174 vs. 152 ± 68 mm^3, t_{22} = 0.32, p = 0.75) or the right hemisphere (164 vs. 167 ± 70 mm^3, t_{22} = 0.04, p = 0.97). In auditory cortex, which overlapped the lesion zone, there was less active cortex in left hemisphere in SJ compared with controls (75 vs. 242 ± 76 mm^3, t_{22} = 2.16, p = 0.04) and no difference in right hemisphere (202 vs. 213 ± 71 mm^3, t_{22} = 0.15, p = 0.88).

Method 2: Amplitude of HDR to McGurk Stimuli

Next, we examined the amplitude of the response to McGurk stimuli within the STS, visual cortex, and auditory cortex ROIs. Because these ROIs were created with independent localizer scans that contained words and not McGurk stimuli, the analysis was not biased (Kriegeskorte et al., 2009, Vul et al., 2009). There was no response in SJ's left STS (0% in SJ vs. 0.11% in controls t_{22}= 4.25, p = 10^{-4}) but the response in SJ's right STS was significantly greater that controls (0.29% in SJ vs 0.13% in controls, t_{71} = 2.57, p = 0.01) (Figure 2.5C). This pattern (less activity than controls in left hemisphere, more activity than controls in right hemisphere) was not found in other ROIs. In visual cortex, there were no significant difference in McGurk amplitude in the left extrastriate cortex (0.07% in SJ vs 0.10% in controls, t_{71} = 0.67, p = 0.50) while right hemisphere showed greater response (0.21% in SJ vs 0.12% in controls, t_{71} = 1.96, p = 0.05). In auditory cortex, SJ's response was significantly weaker in left hemisphere (-0.06% in SJ vs 0.22% in controls, t_{71} = 5.64, p = 3 x 10^{-7}) but was similar to controls in right hemisphere (0.26% in SJ vs 0.19% in controls, t_{71} = 1.33, p = 0.19).

If SJ's right STS subserved new functions because of the lesion to SJ's left STS, we would expect a differential pattern of activity in SJ's right STS compared to other right hemisphere ROIs. To test this idea, we performed an ANOVA on right hemisphere responses to McGurk stimuli across the ROIs between SJ and controls (the variance was computed within subject for SJ and across subjects for controls). A main effect of subject group (SJ vs. controls) would suggest that all right hemisphere ROIs showed different responses between SJ and controls. A main effect of ROI (STS, auditory

cortex, visual cortex) would suggest that a particular ROI was more active, regardless of group. A significant interaction would suggest differential effects between different right hemisphere ROIs between SJ and controls. The ANOVA found a significant interaction between group and ROI ($F_{2,213} = 4.70$, $p = 0.01$) without significant main effects for group or ROI. This suggests that the different ROIs in the right hemisphere responded differently in SJ compared with controls, driven by a greater a response in right STS in SJ compared with controls.

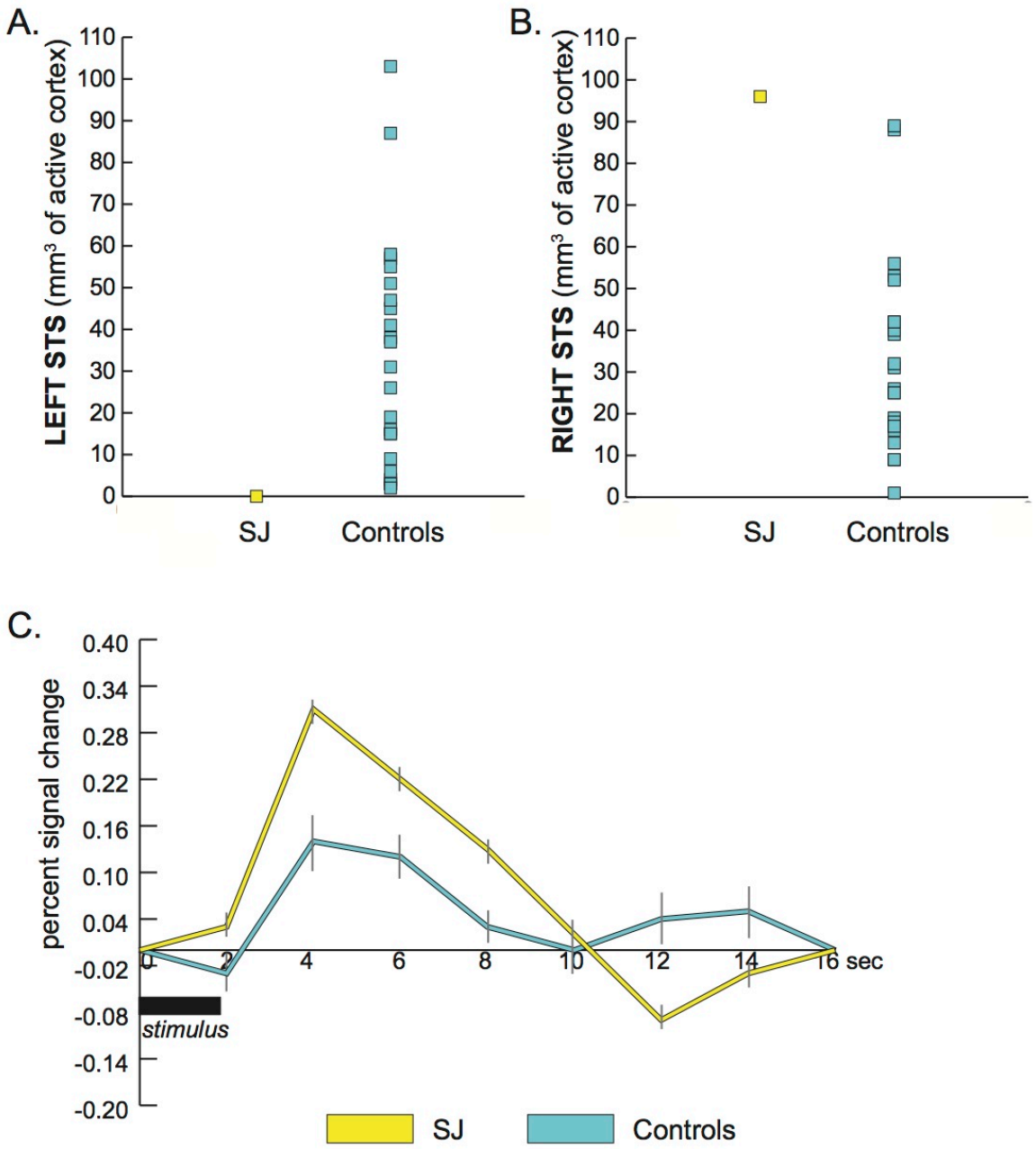

Figure 2.5 Multisensory responses in the STS in SJ and controls
A. Volume of active cortex in the left STS of SJ (yellow) and age-matched controls (blue).
B. Volume of active cortex in the right STS of SJ (yellow) and age-matched controls (blue).
C. Hemodynamic response for SJ (yellow) and healthy controls (blue) in the right STS in response to the McGurk syllable A-“ba”/V-“ga”. Error bars denote standard error of the mean (within-subject variance for SJ and between-subject variance for controls).
Taken from Baum et al. (2012).

Method 3: Whole Brain Analysis

In a third strategy to look for neural differences between SJ and controls, we performed a whole brain analysis of the response to McGurk stimuli. Regions with both increased and decreased responses relative to controls were observed (Table 2.2). The region with the largest area of increased activity in SJ relative to controls was in the right STS. The region with the largest decrease in activity in SJ relative to controls was in the left STS and the remainder of the lesion zone in the left hemisphere.

Increased Activity in SJ

Label	Area (mm^2)
R superior temporal sulcus *(45, -59, 18)*	1250
R superior frontal sulcus *(29, 9, 48)*	1120
L frontomarginal sulcus *(-29, 49, 2)*	935
L central sulcus *(-25, -31, 62)*	748
L angular gyrus *(-45, -69, 22)*	731
R frontomarginal sulcus *(33, 49, -2)*	547

Decreased Activity in SJ

Label	Area (mm^2)
L lateral-posterior temporal, including STS *(-43, 24, 7)*	4449
L postcentral sulcus *(-35, -43, 38)*	695

Table 2.2 Areas of differential activation in SJ and controls
Regions from the whole brain analysis of significant difference in response to McGurk stimuli between SJ and age-matched controls, mapped to the cortical surface. Regions are ranked by area on the cortical surface. Talairach coordinates following anatomical label in *(x, y, z)* format are the weighted center of mass of the cluster.
Taken from Baum et al. (2012).

Amplitude of HDR to Congruent and Non-McGurk Incongruent Stimuli

In addition to McGurk stimuli (which were of greatest interest because they require multisensory integration) we also measured the response to congruent stimuli and non-McGurk incongruent stimuli. In the STS of normal controls, the largest response was to non-McGurk incongruent stimuli with significantly weaker responses to congruent and McGurk stimuli (incongruent stimuli: 0.22% in left STS, 0.25% in right STS compared with congruent: 0.16% in left STS, $t_{22} = 2.74$, p=0.01; 0.17% in right STS, $t_{22} = 3.08$, p=0.01; compared with McGurk: 0.14% in left STS, $t_{22} = 2.41$, p=0.03; 0.14% in right STS, $t_{22} = 3.08$, p=0.01; no significant hemispheric differences) (Figure 2.6A). This response pattern was markedly altered in SJ. Instead of the maximal response to non-McGurk incongruent stimuli observed in controls, SJ had similar amplitudes of response to each stimulus type in her right STS (non-McGurk incongruent = 0.25%, McGurk = 0.29%, congruent = 0.29% , $F_{2,147} = 0.33$, p = 0.72) (Figure 2.6B).

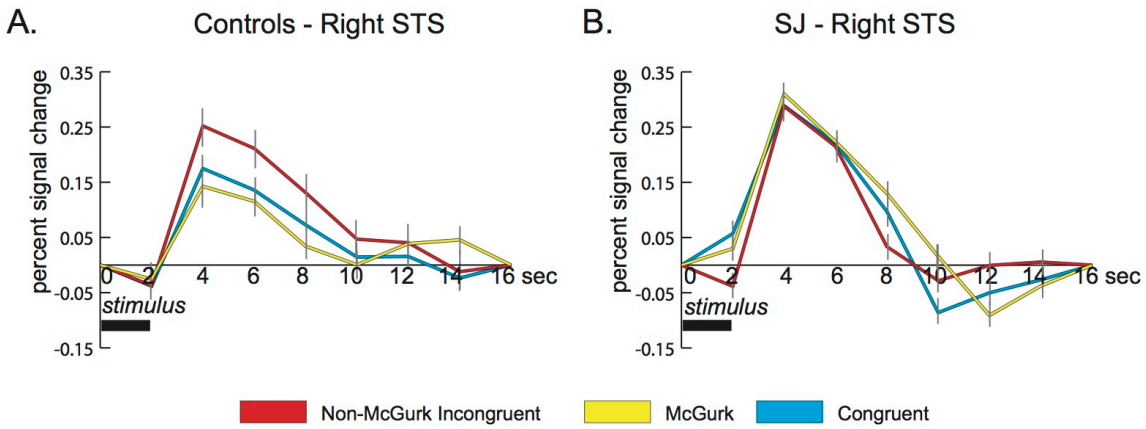

Figure 2.6 Hemodynamic response to all audiovisual stimuli
A. Response to non-McGurk incongruent (red), McGurk (yellow) and congruent (blue) audiovisual stimuli in the right STS of age-matched controls. Error bars denote standard error of the mean across subjects.
B. Response to the same stimuli in the right STS of SJ. Error bars denote standard error of the mean within SJ.
Taken from Baum et al. (2012).

Conclusions

We examined a subject, SJ, whose stroke completely destroyed a large portion of her left temporal lobe, including the left STS. Previous studies have demonstrated a critical role of the left STS in multisensory speech perception (Scott & Johnsrude, 2003, Beauchamp, 2005, Miller & D'Esposito, 2005, Stevenson & James, 2009, Nath & Beauchamp, 2011, 2012). Because temporary disruption of the left STS with TMS impairs multisensory speech perception (Beauchamp et al., 2010) one might expect the lesion suffered by SJ to greatly reduce multisensory integration. Surprisingly, patient SJ showed robust multisensory integration when tested with two independent behavioral tests five years after her stroke.

Evidence suggests that SJ's speech perception abilities changed in the years following her stroke, during which she received extensive rehabilitation therapy. She spent 12 hours a week for approximately 40 weeks a year in the years following her stroke at the Houston Aphasia Recovery Center as well as receiving additional speech and language therapy. SJ and her husband report that this intensive therapy has been extremely beneficial to her recovery. Consistent with this anecdotal report, SJ's speech perception abilities improved following her stroke, from 48% on auditory lexical decision 3 years following the stroke to 87% at 5 years following the stroke (because multisensory integration was only tested 5 years following the stroke, we do not know whether SJ's multisensory abilities showed a parallel improvement.)

Based on the observed improvements in speech perception, neural plasticity and rehabilitation in SJ might have resulted in brain changes, leading to her improved abilities. This would predict different patterns of brain activity during multisensory speech perception in SJ compared with age-matched controls. To test this hypothesis, we studied the neuroanatomical substrates of multisensory speech perception with structural and functional MRI in SJ and 23 age-matched controls. Age-matched controls had large volumes of active multisensory cortex in both the left and right STS when perceiving audiovisual speech. In comparison, speech evoked no activity in SJ's left STS but a larger volume of active cortex in right STS than in any age-matched control. The response amplitude to McGurk stimuli in the right STS was significantly greater than the right STS response in the healthy age-matched controls. These results suggest that SJ's multisensory speech perception may be supported by her right STS. As auditory noise increases, multisensory integration becomes more important (Ross et al., 2007). SJ's diminished auditory abilities immediately following her stroke may have driven the recruitment of right hemisphere areas in the service of multisensory integration for speech comprehension.

A notable finding is that the response amplitude in SJ's right STS to all three types of audiovisual syllables was large and relatively uniform, in contrast with the maximal activation to incongruent stimuli observed in healthy

controls (van Atteveldt et al., 2010, Stevenson et al., 2011). This could reflect an attentional effect, in which healthy subjects automatically process most audiovisual speech, with an enhanced response to incongruent stimuli because they attract attention. SJ's right STS processing of speech may require more conscious effort on her part, resulting in attentional modulation (and enhanced response) for all audiovisual speech stimuli. Indeed, SJ reports that watching speakers on TV (such as a newscast) or conversing with others is especially mentally effortful.

Our results are consistent with a large body of literature showing that the contralesional hemisphere is able to compensate for damage after a brain injury. Left hemisphere strokes often result in aphasia (Dronkers et al., 2004) that resolves (at least partially) over time. Functional imaging studies of these cases have demonstrated increased activity in right-hemisphere homologues of left hemisphere language areas (Buckner et al., 1996, Thomas, 1997, Cao et al., 1999, Blasi et al., 2002, Winhuisen et al., 2005). While these studies used high-level language tasks, such as word retrieval, we observed similar right hemisphere compensation in a low-level task that required integration of auditory and visual speech information.

While the finding that SJ has multisensory integration is surprising based on the McGurk perception literature from healthy controls, it is in line with other reports from aphasics in the literature showing that aphasics are able to integrate sensory information. Champoux et al. (2006) examined a 12 year old child with damage to the right inferior colliculus and noted that when McGurk stimuli were presented in the left hemifield, the patient's perception of the illusion was dramatically reduced. McGurk fusion percepts have also been found in stroke patients whose lesion locations are less well defined (Campbell et al., 1990, Schmid et al., 2009). Youse et al. (2004) describe a patient, JP, who suffered a left hemisphere stroke and perceived the McGurk effect (although poor performance on the auditory-only syllables makes this more difficult to interpret than in SJ). Other audiovisual integration effects have been noted in patients who presented with visual neglect, hemianopia, or both (Frassinetti et al., 2005). An important distinction is between auditory-visual language stimuli in which both modalities are presented in their natural speech form (i.e. auditory "ba" + video of speaker saying "ba") with an orthographic representation (i.e. auditory "ba" + printed letters "ba"). Although orthographic auditory-visual tasks also recruit the STS (Raij et al., 2000, van Atteveldt et al., 2004, Blau et al., 2008) there are differences between letter-speech and audiovisual speech processing (Froyen et al., 2010) and lesions might be expected to differentially impair these two tasks. For instance, Hickok et al. (2011) found that Broca's aphasics were impaired on an auditory-visual grapheme discrimination task.

We observed significant variability within our population of 23 age-matched controls, which may be linked to individual differences in

multisensory integration and language ability (Kherif et al., 2009, Nath et al., 2011, McGettigan et al., 2012, Nath & Beauchamp, 2012). Because we do not have pre-injury data for SJ, we cannot refute the null hypothesis that her right hemisphere subserved multisensory integration even before the stroke and that no cortical reorganization occurred. However, the observation that SJ's volume of speech-evoked activity in right STS was greater than in any age-matched control (and that no activity was observed in SJ's left STS, far less than in any age-matched control) supports a neural plasticity explanation. SJ's extensive rehabilitation efforts are similar to those known to cause dramatic reorganization in language networks, such as in illiterate adults undergoing literacy training (Carreiras et al., 2009).

While our study does not provide direct evidence that the activity observed in SJ's right STS is critical for her multisensory abilities, other studies have shown that disrupting the right hemisphere of recovered aphasia patients using TMS (Winhuisen et al., 2005), intracarotid amobarbital (Kinsbourne, 1971, Czopf, 1979) or even additional infarcts (Turkeltaub et al., 2011) results in profound language impairments. We hypothesize that a similar manipulation, such as TMS of SJ's right STS, would greatly reduce her multisensory speech perception.

Taken from Baum, S.H., Martin, R.C., Hamilton, A.C., and M.S. Beauchamp (2012) "Multisensory speech perception without the left superior temporal sulcus." NeuroImage 62(3) 1825-1836.

CHAPTER 3: INCREASED NEURAL VARIABILITY IN MULTISENSORY SPEECH PERCEPTION IN OLDER ADULTS

Introduction

The ability of older adults to understand both auditory-only and audiovisual speech declines with age (Sumby & Pollack, 1954, Dubno et al., 1984, Grant & Seitz, 2000, Gosselin & Gagne, 2011). This decline extends to other important cognitive functions, such as memory, visuospatial abilities, and speed of information processing (Cerella & Hale, 1994, Jenkins et al., 2000, Hedden & Gabrieli, 2004, Peich et al., 2013). Interestingly, performance declines with age are not uniform across multiple trials of the same task. Older adults exhibit much greater variability in performance: on some trials, older adults perform as well as younger adults, but on other trials, older adults perform much worse (Lovden et al., 2007, Bielak et al., 2013, Vandermorris et al., 2013). This type of performance decline, referred to as increased intrasubject variability, may be a particularly sensitive measure of age-related cognitive changes (Butts & Goldman, 2006).

We hypothesized that the increased intrasubject variability observed behaviorally should have a neural counterpart. That is, across multiple presentations of the same stimulus, the neural response evoked by the stimulus should show greater variability in older subjects than younger subjects. Increased neural variability with age has been observed in the visual cortex of experimental animals (rhesus macaques and cats) presented with simple visual stimuli (Schmolesky et al., 2000, Hua et al., 2006, Liang et al., 2010).

To determine if the same effect is found in human subjects perceiving complex audiovisual speech, we used rapid event-related blood-oxygen level

dependent magnetic resonance imaging (BOLD fMRI). Specifically, we predicted that repeated presentations of identical speech stimuli would evoke fMRI responses that were more variable in older adults than in younger adults. We examined neural responses to audiovisual speech consisting of single syllables using two complementary methods. First, we used a region-of-interest (ROI) analysis focused on the three core areas of the multisensory speech perception network: auditory cortex, visual cortex, and the superior temporal sulcus (STS). Second, we used a voxel-wise whole-brain analysis.

Methods

Subjects

All subjects provided informed consent and were compensated for their time in accordance with an experimental protocol approved by the Committee for the Protection of Human Subjects of the University of Texas Health Science Center at Houston. 24 healthy older adults participated in the study. Five subjects were excluded (see section: *Behavioral tests for older adults and exclusion criteria*) leaving 19 subjects whose data are reported here (53-70 years, 12 female, mean age 63.0 years, 17 right-handed and 2 ambidextrous). Data from the ambidextrous subjects was similar to that of the right-handed subjects so they were analyzed together. The young adult cohort consisted of 14 subjects (14 subjects, 20-39 years, 6 female, mean age 26.1 years, 14 right-handed). Handedness was assessed by the Edinburgh Handedness Inventory (Oldfield, 1971).

Overview of fMRI experiment and analysis

We used two independent methods for fMRI analysis: region-of-interest (ROI) and voxel-wise whole-brain analysis, which give complementary information about brain activity (Friston et al., 2006, Saxe et al., 2006). ROI analysis allows us to examine areas for which we have an *a priori* hypothesis and does not require that data be transformed to a brain template, thus allowing for differences in individual anatomy. Furthermore, it limits Type I errors by limiting the number of statistical tests to a handful of ROIs (Poldrack, 2007). However, in a voxel-wise whole-brain analysis there are no possible biases in how the ROIs

are defined and potentially interesting patterns of results are not missed by averaging over a group of defined voxels.

Block-design localizer

A block-design localizer was used to generate ROIs. Each block contained 10 two-second trials, one word per trial, followed by 10 seconds of fixation baseline. Each trial contained a single word from a bank of digital video recordings of 105 single-syllable words (*e.g.* “view”, “door”, “make”) spoken by a female native English speaker. Words were selected from the MRC Psycholinguistic Database (Wilson, 1988). Auditory-only words consisted of the auditory component of each video with a white visual fixation crosshairs and visual-only words consisted of only the visual component of the video recording.

In older adults, the localizer scan series contained six blocks (two auditory-only, two visual-only and two audiovisual blocks in random order). Each block contained a target trial (the word “press”) of the same type (auditory-only, visual-only, or audiovisual) as the other stimuli in the block; subjects were instructed to pay attention to each stimulus and press a response button only during target trials. In younger adults, ten blocks were presented (five auditory-only and five visual-only in random order) with no target trials.

fMRI responses to audiovisual speech syllables

For the main experiment, stimuli were presented in two-second trials in a rapid event-related design. Each trial contained a single audiovisual syllable, consisting of McGurk (auditory “ba” + visual “ga”, auditory “pa” + visual “ka”), non-McGurk incongruent (auditory “ga” + visual “ba”, auditory “ka” + visual “pa”),

congruent (“ba”, “ga”, “da”, “pa”, “ka” and “ta”), target (audiovisual “press” in older adults, audiovisual “ma” in younger subjects) and fixation trials (fixation crosshairs only). Subjects were instructed to respond with a button press only to target trials and to make no response to all other trials. Behavioral data in the scanner was not collected for two younger subjects. Nearly all subjects performed very well on this task (18/19 older adults at 100% accuracy; 10/12 younger adults at 100% accuracy) suggesting a high degree of alertness (no significant difference between groups, $t_{29} = 0.6$, $p = 0.57$).

Because we used target trials in the localizer scan in addition to the main experiment in older adults, the designated target trial was changed so that it was the same type of stimulus (auditory-only, visual-only, or audiovisual word) as the rest of the stimuli within that block. This change in target trial was extended through the rapid event related scan series for consistency. For both older and younger adults, the target trials in the rapid event related runs were analyzed with a separate regressor in the generalized linear model, which allowed us look at the activation elicited by the target trials independently from the audiovisual syllables.

Fixation trials were used as the baseline for the analysis. Target trials were found to evoke brain responses related to motor planning and execution (Beauchamp et al., 2007). The brain response to all audiovisual syllables (McGurk, non-McGurk incongruent, and congruent) was similar, so they were combined for further analysis and only the average across stimulus types (excluding target trials) is reported.

The total length of each video was cropped with digital video editing software (iMovie, Apple Computer) such that each clip started and ended in a neutral, mouth-closed position. Each video stimulus varied in length from 1.7 to 1.8 seconds followed by fixation crosshairs for the remainder of the trial (the crosshairs were always presented in the same screen location as the mouth of the talker visible during other trials in order to minimize eye movements). Prior to the scan, a volume check was conducted for each subject outside the scanner without the presence of scanner noise. Sample videos from the experiment were played and the volume was adjusted so that the volume was “as loud as possible without being uncomfortable or hurting in any way”. After each scan series subjects were asked if they could hear the stimuli presented and if any volume adjustments were necessary.

Each scan series contained multiple trials. The number of trials in each scan series was as follows: older subjects (*n* = 6): 75 audiovisual syllables, 25 fixation trials, 20 target trials; older subjects (*n* = 13): 80/25/15; younger subjects (*n* = 5): 110/35/10; younger subjects (*n* = 9): 100/40/10. Two scan series were collected in each older subject. Three or four scan series were collected in each younger subject, but only the first two were analyzed to roughly equate the amount of data collected in younger and older subjects.

MRI and fMRI analysis

Two T1-weighted MP-RAGE anatomical MRI scans were collected at the beginning of each scanning session with a 3 Tesla whole-body MR scanner (Phillips Medical Systems). The two anatomical scans were aligned to each other

and averaged in order to provide maximal gray-white matter contrast. These scans were then used to create a cortical surface model using FreeSurfer (Dale et al., 1999, Fischl et al., 1999) for visualization in SUMA (Argall et al., 2006). For the fMRI scan series, T2* weighed images were collected using gradient echo-planar imaging (TR = 2000 ms, TE = 30 ms, flip angle = 90°) with in-plane resolution of 2.75 x 2.75 mm. Auditory stimuli were presented through MRI-compatible in-ear headphones (Sensimetrics, Malden, MA) which were covered with ear muffs to reduce the amount of noise from the scanner. Visual stimuli subtending approximately 20 x 30 degrees of visual angle were presented on a projection screen with an LCD projector and viewed through a mirror attached to the head coil. Responses to the target trials were collected using a fiber-optic button response pad (Current Designs, Haverford, PA). Analysis of the functional scan series was conducted using Analysis of Functional NeuroImages (AFNI) (Cox, 1996).

fMRI analysis: response amplitude and variability

The voxel-wise analysis was carried out with the AFNI function *3dDeconvolve,* which uses maximum-likelihood estimation in the context of the generalized linear model (GLM). TENTzero functions were used to estimate the individual hemodynamic response function (using the option *–iresp*) and standard deviation of each response function (using the option *–sresp*) in each voxel for each stimulus type, beginning at stimulus onset and ending 16 seconds later for single syllables and 26 seconds later for blocks of words. For single syllables, we estimated the amplitude of the response as the mean of the response at 4

seconds and 6 seconds after stimulus onset (the peak of the hemodynamic response function). To estimate BOLD variability within each subject for single syllables, the standard deviation at the 4-second and 6-second time points of each impulse response function were averaged to produce a single value per voxel.

An important methodological point is that the standard deviation was calculated from the response in each individual voxel, not from a fixed hemodynamic response function shape, such as a gamma variate. The use of a fixed function could introduce a confound because of differences in the shape of individual subjects' hemodynamic response functions. For instance, if older people had slightly broader hemodynamic response functions, then their deviation from a fixed function would be greater, unrelated to trial-to-trial variability in the amplitude of the response.

Region-of-interest selection

Data were first analyzed for each subject individually in native image space. ROIs were selected to target brain areas that are reliably active during multisensory speech perception (Nath & Beauchamp, 2011). A combination of anatomical and functional criteria was used. The anatomic parcellation of the cortical surface was constructed from each individual subject's structural scans with FreeSurfer (Fischl et al., 2004, Destrieux et al., 2010). Functional criteria were constructed from the independent localizer runs, eliminating bias (Kriegeskorte et al., 2009).

We considered three contrasts when constructing the three ROIs: auditory

words vs. fixation baseline, visual words vs. fixation baseline, and audiovisual words vs. fixation baseline. The STS ROI was defined by finding all voxels in the posterior half of the anatomically parcellated STS that showed a significant response ($t > 2$ for auditory-only word blocks *vs.* baseline and $t > 2$ for visual-only word blocks *vs.* baseline). For 5 out of 19 older adults, no voxels in the left STS met this criterion, so an alternative criterion was used ($t > 2$ for audiovisual word blocks *vs.* baseline). Spread of neural activity often observed in healthy older adults might have resulted in the 'spill' of much of the multisensory activity outside the bounds of the posterior STS for these subjects, and because of this the robust multisensory activation required to meet the threshold of the conjunction definition for the STS ROI may not have been met. The auditory cortex ROI was defined by finding voxels in the anatomically parcellated transverse temporal gyrus, lateral superior temporal gyrus and planum temporale that were significantly active during the auditory-only blocks ($t > 2$ for auditory-only word blocks *vs.* baseline). The extrastriate visual cortex ROI was defined by finding voxels in the anatomically parcellated extrastriate lateral occipitotemporal cortex that were active during the visual-only blocks ($t > 2$ for visual-only word blocks *vs.* baseline).

Whole-brain analysis

For the whole-brain voxel-wise analysis, subjects' individual data were first aligned to the N27 atlas brain (Mazziotta et al., 2001) using the AFNI function *auto_tlrc*. Blurring kernels of approximately 3-6 mm have been found to be the most sensitive for detecting activation clusters (Skudlarski et al., 1999). We

chose a 3 x 3 x 3 mm FWHM Gaussian kernel to minimize blurring between adjacent ROIs.

To conduct a voxel-wise search for any differences in response amplitude, the average response amplitude (average of the response to all non-target audiovisual speech stimuli relative to fixation baseline at the 4 and 6 second time points) was calculated in each voxel in each subject. *3dttest++* was used to perform an unpaired t-test for every voxel in standard space between the old and young adult groups. The results were mapped from the MRI volume to the cortical surface with *3dSurf2Vol* and masked with the group t-statistic ($t > 2$ for the contrast of all audiovisual syllables *vs.* baseline). After the voxel-wise t-test we preformed a clustering technique (Xiong et al., 1995). This finds only voxels that are significantly active above a particular threshold and spatially contiguous. The probability of finding two voxels above a particular threshold *and* being adjacent is much smaller than the chance of a single voxel above that threshold (Forman et al., 1995). Using the AFNI program *slow_surf_clustsim.py*, we estimated that a cluster with a size of 160mm^2 would have a corrected p-value of 0.045. A clusterizing filter on the surface (*SurfClust)* was applied and only regions larger than 160 mm^2 (and $t > 2$ for the contrast of all audiovisual syllables *vs.* baseline) are reported.

To conduct a voxel-wise search for differences in intersubject variability, the MATLAB function *vartestn* was used to perform a Bartlett's multiple sample test for equal variances on the response amplitudes, followed by clusterizing.

To conduct a voxel-wise search for differences in intrasubject variability,

an unpaired t-test between groups was performed on the standard deviation of the response at each voxel (*3dttest++*) followed by clusterizing.

Motion correction

Functional data for each subject was first aligned to the averaged anatomical dataset for that subject and then motion-corrected using local Pearson correlation with the AFNI script *align_epi_anat.py* (Saad et al., 2009). For each volume, an estimate of the amount of motion in each direction, relative to the reference, was produced. These estimates were used as regressors of no interest in the fMRI analysis. To capture a single value describing the amount of head motion in each subject, the standard deviation of each motion direction across time was averaged across motion directions.

Because differences in head movements during fMRI may confound intergroup comparisons (Van Dijk et al., 2012, Wylie et al., 2012), two older subjects were excluded for large head motions (standard deviation of motion regressor > 3mm). Head movements tend to be larger in healthy aging and patient populations.

We also performed a “motion scrubbing” procedure developed by Power and colleagues (Power et al., 2012). First, motion estimates at each time point were calculated in each of the six motion directions (rotational measures: roll, pitch, yaw, and displacement measures: superior, left, and posterior directions). Rotational displacement measures were converted to millimeters using the formula:

$$d = R*(\pi/180)*r$$

where R is the rotation in degrees and r is the radius (we used r = 50 mm as prescribed by Power). To express the total amount of motion for each time point in a single value, the absolute value of the displacement in each direction was summed, where the total displacement at the *i*th data point was:

$$D_i = |d_\alpha| + |d_\beta| + |d_\gamma| + |d_x| + |d_y| + |d_z|.$$

Then the framewise displacement for the *i*th time point was calculated as:

$$FD_i = D_{(i-1)} - D_i$$

to express instantaneous head motion. The scrubbing threshold was half of the smallest voxel dimension, as recommended by Power (EPI volumes were collected using a 2.75mm isotropic voxel, therefore we used a threshold of 1.375 mm). The GLM analysis was then completed a second time, excluding the data points with a FD exceeding this threshold.

Behavioral tests for older adults and exclusion criteria

Tests for vision, hearing, and cognitive function were conducted on the older adults. In total, 5 subjects were excluded from the fMRI analysis (two for impaired hearing, three for fMRI data quality concerns). Of the 24 total subjects recruited for the study, only the data from the remaining 19 subjects is reported here.

Vision was assessed using a Snellen eye chart at the same visual ability they would have in the scanner (*i.e.* corrective contacts, if worn, but not glasses). Each eye was tested separately. The range of acuities was 20/20 to 20/70.

Hearing was evaluated using a modified Bekesy threshold test at 500 and 2000 Hz (Price, 1963). No subjects were excluded for poor simple hearing. The

range of hearing thresholds with both ears for normal aging in the age range of the subjects in our study is 10.5±11.1 dB – 14.3±12.6 dB at 500 Hz and 14.1±15.3 dB – 28.6±20.9 dB at 2000 Hz (Brant & Fozard, 1990). Our subjects had a range of 6.7 dB – 30.9 dB (mean: 13.5 dB, standard deviation: 4.8 dB) for 500 Hz and 7.3 – 39.4 dB (mean: 15.8 dB, standard deviation: 8.1 dB) for 2000 Hz. Therefore, our subjects were within the normal range (within two standard deviations of the mean) for hearing thresholds based on their age. However, speech abilities decline at a different rate than pure audiometric measures later in life (Divenyi et al., 2005), therefore we also tested identification of auditory-only and audiovisual syllables. Most subjects scored near ceiling on auditory-only syllable identification (83% - 100%, average performance 93%) and audiovisual syllables. Two subjects scored poorly on auditory-only syllable identification (<70%) and were excluded from the analysis. One older subject was excluded because an initial analysis found response amplitudes more than 3 standard deviations greater than the mean.

Cognitive function was assessed using the standardized Mini Mental State Examination (MMSE) (Folstein et al., 1975). All subjects' scores indicated no decline in cognitive function (scores ranged from 26-30, mean MMSE 28.4, scores 25 out of 30 points or greater indicate normal cognitive function).

Results

Responses in the ROIs to audiovisual speech

Our initial analysis focused on three ROIs implicated as critical nodes in the network for multisensory speech perception: the left superior temporal sulcus (STS), the left auditory cortex, and the left extrastriate visual cortex.

Mean and standard deviation of the hemodynamic response across subjects

The left STS showed a robust hemodynamic response to audiovisual syllables that was similar in older and younger subjects (Figure 3.1A). An unpaired t-test with percent signal change in the left STS as the dependent measure revealed slightly greater amplitude of mean response in younger adults (0.19% in younger adults *vs.* 0.12% in older adults, $t_{31} = 2.1$, $p = 0.048$). There were no significant differences in auditory cortex (0.26% *vs.* 0.24%, $t_{31} = 0.5$, $p = 0.63$) or visual cortex (0.16% *vs.* 0.10%, $t_{31} = 1.5$, $p = 0.15$).

The standard deviation of the response across subjects was similar between old and young (Figure 3.1B; left STS: SD of 0.08% for younger adults *vs.* 0.12% for older adults, Bartlett's multiple sample test for equal variances $\chi_1^2 = 2.5$, $p = 0.12$; left auditory cortex: 0.14% *vs.* 0.12%, $\chi_1^2 = 0.7$, $p = 0.41$; left visual cortex: 0.08% *vs.* 0.12%, $\chi_1^2 = 2.3$, $p = 0.13$), indicating that the older adult group did not show greater intersubject (across subject) variability across the three *a priori* ROIs.

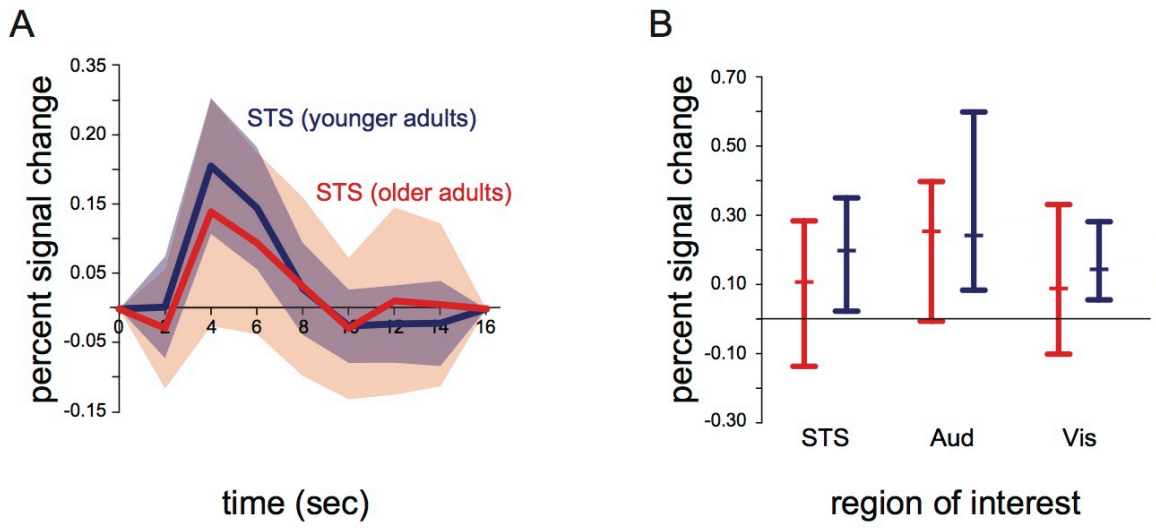

Figure 3.1 BOLD responses to audiovisual speech
A: Average hemodynamic response to audiovisual syllables in the left STS for older adults (red) and younger adults (blue). Shaded region indicates standard deviation of the group response (intersubject variability).
B. Response amplitudes in the left STS (STS), left auditory cortex (Aud), and left visual cortex (Vis) across all older adults (red) and younger adults (blue). Error bars show the complete range of data (subjects with maximum and minimum response); middle bar shows subject with median response.

Variability of the hemodynamic response within subjects

An important behavioral difference between older and younger subjects is the variability across trials *within* individual subjects (intrasubject variability). To measure a neural parallel of this value, we first calculated the hemodynamic response function in every voxel and then measured the standard deviation from this response across trials.

Figures 3.2A and 3.2B show data from the left STS of a representative older and younger subject, respectively (selected by choosing the individuals whose standard deviation was closest to the mean standard deviation for the group). While the response amplitudes are similar for the two subjects, the variability at each time is much larger in the older subject.

To quantify this difference, we averaged the standard deviation from the peak of the response (4 and 6 seconds after stimulus onset) to produce a single number for intrasubject variability for each ROI for each subject, which was then plotted (Figure 3.2C-E). For each ROI, there was much greater within-subject standard deviation in older subjects (STS: 0.14% in older adults *vs.* 0.09% in younger adults, $t_{31} = 4.2$, $p = 2 \times 10^{-4}$; auditory cortex: 0.18% *vs.* 0.11%, $t_{31} = 5.9$, $p = 10^{-6}$; visual cortex: 0.18% *vs.* 0.08%, $t_{31} = 7.0$, $p = 10^{-8}$).

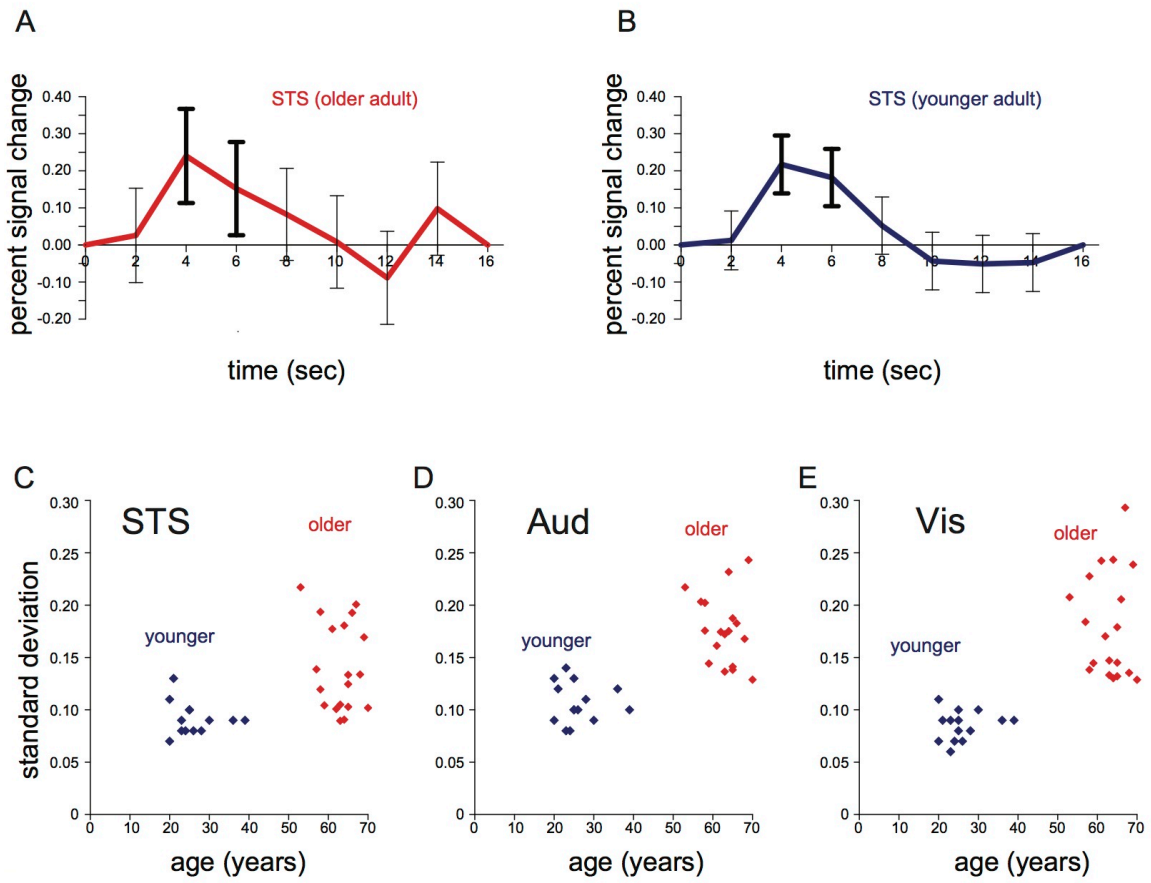

Figure 3.2 Intrasubject variability in older and younger subjects
A: Hemodynamic response in the left STS of a single older adult (subject JI). Error bars indicate standard deviation of the response within that subject (intrasubject variability) at each time point. The variability at the 4 second and 6 second time points (bold error bars) were used for group analysis.
B: Hemodynamic response in the left STS of a single younger adult (subject HU).
C: Scatter plot of age vs. within-subject standard deviation of the STS response. Each blue symbol represents a single younger adult; each red symbol a single older adult. The lines show the mean of the within-subject standard deviation across each group. The brackets show the results of an unpaired t-test between the within-subject standard deviation in each group.
D: Scatter plot of age vs. within-subject standard deviation of the left auditory cortex response.
E: Scatter plot of age vs. within-subject standard deviation of the left visual cortex response.

Whole-brain Analysis

Our first set of analyses was limited to our three *a priori* ROIs created using block-design localizers. To overcome this limitation, and to prevent any biases introduced by slight differences in the localizers between old and young subjects, our second set of analyses examined the entire brain. First, we selected all voxels that showed a significant positive response ($t > 2$ for all audiovisual syllables *vs.* baseline) to audiovisual syllables across old and young subjects (Table 3.1 and Figure 3.3A).

Label	**Area (mm^2)**	**Peak t-value**
R fusiform gyrus, inferior occipital gyrus, middle occipital gyrus *(26, -71, -6)*	4593	6.2
R fusiform gyrus, inferior occipital gyrus, middle occipital gyrus *(55, -21, 2)*	3054	4.3
R superior temporal sulcus and gyrus *(55, -21, 2)*	2361	7.4
L superior temporal gyrus *(-46, -33, 16)*	2315	5.8
L middle occipital gyrus *(-23, -93, 5)*	560	4.9
R supramarginal gyrus and subcentral gyrus *(49, -6, 44)*	219	3.9
Total	13552	

Table 3.1 Activation to audiovisual speech syllables across both younger and older adults.
Regions are ranked by area (only clusters greater than 160 mm^2 are reported) and Talairach coordinates following anatomical label in *(x, y, z)* format are the weighted center of mass of the cluster.

Mean and standard deviation of the hemodynamic response across subjects

We compared the mean amplitude of response between young and older adults. As shown in Table 3.2 and Figure 3.3B, we found regions in right and left visual cortex and right superior temporal sulcus (total area on the cortical surface = 1000 mm^2) where younger adults had greater response amplitude than older adults (peak t-statistic = 4.3, $p = 2 \times 10^{-5}$). There were no regions where response amplitude was greater in older subjects.

Increased Response Amplitude in Younger Adults

Label	**Area (mm^2)**	**Peak t-value**
R occipital pole *(13, -94, 5)*	302	3.3
R superior temporal sulcus *(47, -36, 4)*	273	3.8
L occipital pole *(-17, -93, 5)*	223	4.3
L subcentral sulcus *(-53, -21, 13)*	202	3.6
Total	1000	

Table 3.2 Whole-brain response amplitude in younger and older adults. Regions are ranked by area (only clusters greater than 160 mm^2 are reported) and Talairach coordinates following anatomical label in *(x, y, z)* format are the weighted center of mass of the cluster.

Next, we compared the standard deviation of the response at every brain voxel. As shown in Table 3.3 and Figure 3.3C, the amplitude of response was more variable across older subjects in bilateral visual cortex and right superior temporal cortex (total area = 4367 mm^2, peak Bartlett's χ_1^2= 29, p = 9 x 10^{-8}). There were no regions where the variability across subjects was greater in younger adults.

Increased Intersubject Variability in Older Adults

Label	**Area (mm^2)**	**Peak χ^2**
L inferior occipital gyrus *(-30, -75, -12)*	1819	29
R inferior occipital gyrus *(24, -81, -8)*	1552	27
R superior temporal gyrus *(62, -8, 2)*	438	15
R transverse temporal gyrus *(44, -24, 10)*	200	7.8
R fusiform gyrus *(34, -49, -19)*	194	9.7
L superior frontal gyrus *(-1, 0, 56)*	164	27
Total	4367	

Table 3.3 Whole-brain intersubject variability in younger and older adults. Regions are ranked by area (only clusters greater than 160 mm^2 are reported) and Talairach coordinates following anatomical label in *(x, y, z)* format are the weighted center of mass of the cluster.

Variability of the hemodynamic response within subjects

We calculated the variability of the response *within* each subject at each voxel, and then compared the two groups. Many brain regions showed greater intrasubject variability in older adults (Table 3.4 and Figure 3.3D), including left auditory cortex and bilateral extrastriate visual cortex (total area = 7026 mm^2; peak t-statistic = 5.8, $p = 2 \times 10^{-6}$). There were no regions where intrasubject variability was greater in younger adults.

Increased Intrasubject Variability in Older Adults

Label	**Area (mm^2)**	**Peak t-value**
R superior temporal gyrus *(54, -19, 1)*	2341	5.0
L middle and inferior occipital gyri *(-38, -69, -3)*	1943	5.8
L planum temporale *(-47, -29, 13)*	1697	4.9
R inferior occipital gyrus *(37, -61, -13)*	882	4.2
R superior temporal sulcus *(45, -56, 5)*	163	5.3
Total	7026	

Table 3.4 Whole-brain intrasubject variability in younger and older adults. Regions are ranked by area (only clusters greater than 160 mm^2 are reported) and Talairach coordinates following anatomical label in *(x, y, z)* format are the weighted center of mass of the cluster.

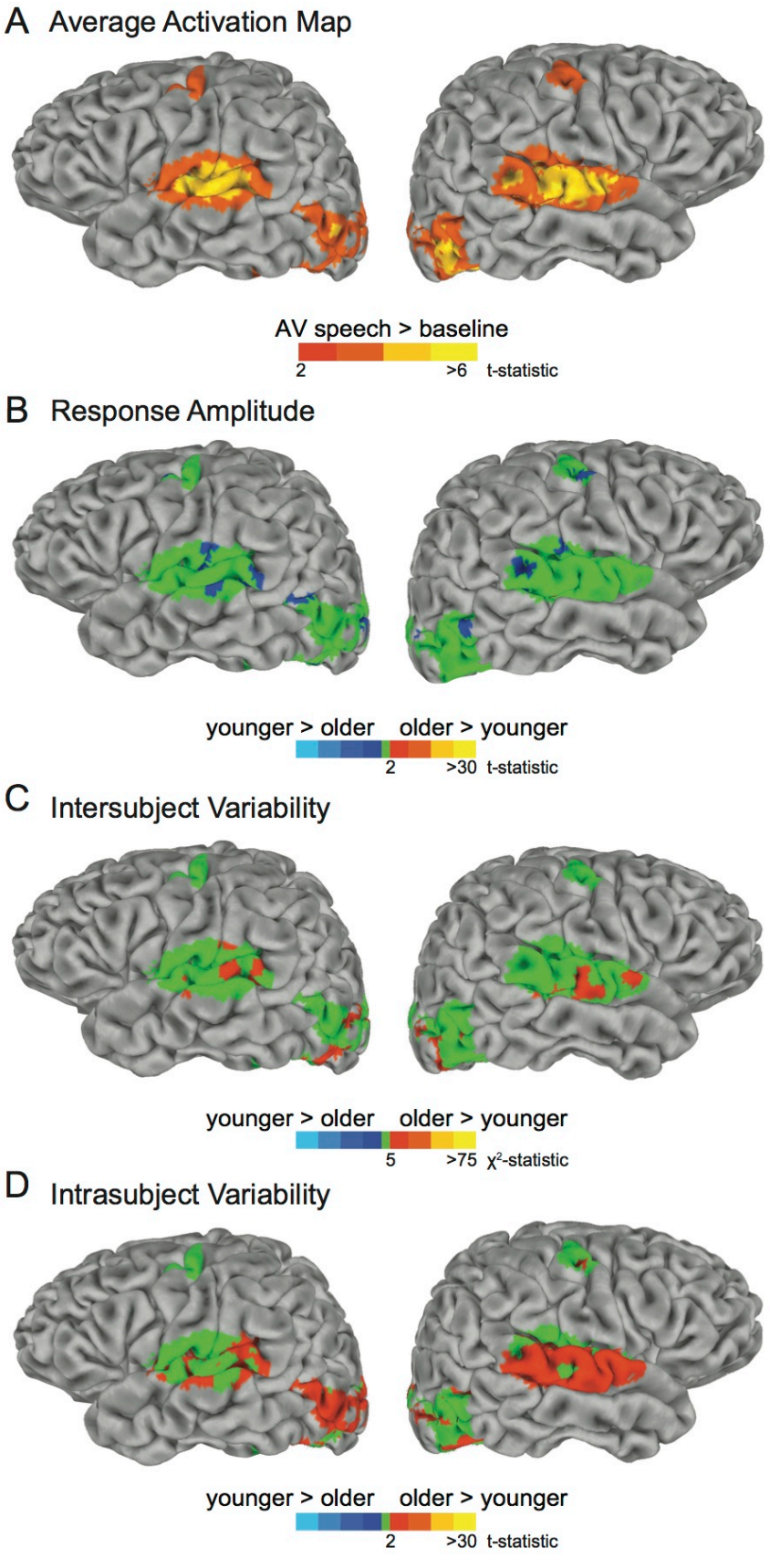
A Average Activation Map
AV speech > baseline
2
>6
t-statistic
B Response Amplitude
younger > older
older > younger
2
>30
t-statistic
C Intersubject Variability
younger > older
older > younger
5
>75
χ²-statistic
D Intrasubject Variability
younger > older
older > younger
2
>30
t-statistic

Figure 3.3 Whole-brain analysis of differences in older and younger adults
Whole-brain group analysis results are displayed on lateral views of the left and right hemisphere of the cortical surface model of the N27 atlas brain.
A: Regions that show a significant positive response ($t > 2$ for all audiovisual syllables *vs.* baseline) to audiovisual speech in both older and younger adults.
B: Differences in response amplitude to audiovisual speech in active regions from **(A)**. Green regions indicate no difference in response amplitude; blue regions indicate areas of greater response amplitude in younger adults.
C: Differences in intersubject variability (variability of the amplitude of the BOLD response *across* subjects). Orange regions indicate areas with greater response variability in older adults.
D: Differences in intrasubject variability (variability of the amplitude of the BOLD response *within* each subject). Orange regions indicate areas with greater intrasubject variability in older adults.

Potential confound: differences in head movements between younger and older adults

Head movements can confound intergroup comparisons (Power et al., 2012, Van Dijk et al., 2012). While our estimates of head movements were small, they were larger in older than younger adults (0.48 mm *vs.* 0.32 mm, $p = 0.004$). In addition to the standard analysis techniques of motion correction and inclusion of motion estimates in regressors of no interest in the general linear model used to reduce the effects of head motion of fMRI data, we performed a number of additional analyses. First, we used the "motion scrubbing" procedure developed by Power and colleagues (Power et al., 2012). Older adults had slightly greater number of data frames flagged for removal (average of 6.5 data frames in older adults *vs.* 2.2 data frames in younger adults, $t_{31} = 4.1$, $p = 3 \times 10^{-4}$). Our finding of increased BOLD signal variability in older adults remained unchanged after the scrubbing procedure (scrubbed data: STS: 0.13% *vs.* 0.09%, $t_{31} = 3.4$, $p = 0.002$; auditory cortex: 0.18% *vs.* 0.10%, $t_{31} = 5.3$, $p = 8 \times 10^{-6}$; visual cortex: 0.17% *vs.* 0.09%, $t_{31} = 5.5$, $p = 1.6 \times 10^{-5}$).

There was no correlation in both older and younger adults between amount of head motion and variability (older adults: $r = 0.04$, $p = 0.88$; younger adults: $r = 0.16$, $p = 0.59$) (Figure 3.4). An ANCOVA was performed the ROI data with age group as one factor, head motion as a covariate, and standard deviation of the fMRI response as the dependent variable and revealed no interaction between age group and head motion in any ROI ($p > 0.68$), and the finding of increased intrasubject variability in older adults remained significant

(STS: $F_{1,29} = 13$, $p = 0.001$; auditory cortex: $F_{1,29} = 25$, $p = 3 \times 10^{-5}$; visual cortex: $F_{1,29} = 34$, $p = 2 \times 10^{-6}$). A whole-brain ANCOVA that included the amount of head motion in each subject as a covariate gave results nearly identical to the analysis without the head motion covariate.

Finally, we discarded the six older adults with the greatest amount of head motion (resulting in a total group size of 13 older adults), which rendered the differences in head movements between groups insignificant (0.39 mm *vs*. 0.32 mm, $p = 0.10$), but left the main finding of increased intrasubject variability in older subjects intact (left STS: 0.14% *vs.* 0.09%, $t_{25} = 3.7$, $p = 0.001$; left auditory cortex: 0.18% *vs.* 0.11%, $t_{25} = 5.4$, $p = 10^{-4}$; visual cortex: 0.17% *vs.* 0.08%, $t_{25} = 6.7$, $p = 5 \times 10^{-7}$).

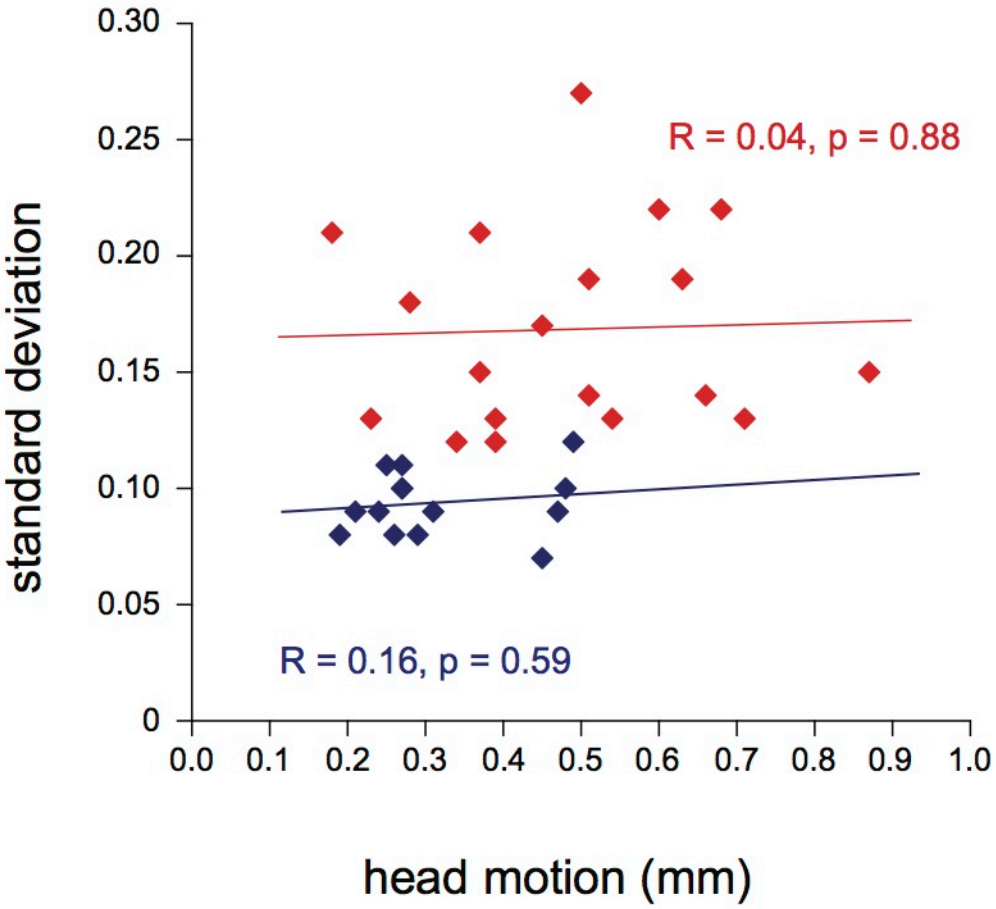

Figure 3.4 Relationship between head motion and variability
Scatter plot of average head motion vs. the average standard deviation of the response to audiovisual syllables in each older subject (red) and each younger subject (blue).

Behavioral performance

Before the scanning session, subjects were tested on a variety of behavioral tests, including auditory, visual, and cognitive measures. We found no correlation between test of auditory thresholds and the amplitude of response (STS: r = -0.16, p = 0.51 at 500 Hz, r = 0.21, p = 0.39 at 2000 Hz; auditory cortex: r = -0.23, p = 0.34 at 500 Hz, r = -0.33, p = 0.17 at 2000 Hz; visual cortex: r = -0.18, p = 0.46 at 500 Hz, r = 0.15, p = 0.54 at 2000 Hz) or the variability of the response (STS: r = 0.03, p = 0.90 at 500 Hz, r = -0.14, p = 0.57 at 2000 Hz; auditory cortex: r = 0.23, p = 0.34 at 500 Hz, r = -0.04, p = 0.87 at 2000 Hz; visual cortex: r = 0.12, p = 0.62 at 500 Hz, r = -0.11, p = 0.65 at 2000 Hz).

Similarly, we found no correlation between visual acuity and the amplitude of response (STS: r = -0.23, p = 0.34; auditory cortex: r = -0.05 p = 0.84; visual cortex: r = -0.16, p = 0.51) or the variability of the response (STS: r = 0.11, p = 0.65; auditory cortex: r = -0.06, p = 0.81; visual cortex: r = 0.15, p = 0.54), nor any correlation between MMSE score and the amplitude of response (STS: r = -0.36, p = 0.13; auditory cortex: r = 0.04, p = 0.87; visual cortex: r = 0.002, p = 0.99) or the variability of the response (STS: r = -0.11, p = 0.65; auditory cortex: r = -0.31, p = 0.20; visual cortex: r = -0.23, p = 0.34).

Performance on the syllable identification task was near-ceiling (auditory-only syllables: 93%, audiovisual syllables: 98%). Because of these near-ceiling values, no correlation with the fMRI data was performed.

Conclusions

We compared brain responses to audiovisual speech in healthy older and younger adults using fMRI. The most important finding was greater intrasubject variability in the older adults: across multiple trials of the same stimulus, older adults had greater variability in their brain responses than younger adults. This was true across all of the brain areas that responded to audiovisual speech and was confirmed with two independent types of analysis (ROI and whole-brain). We also observed two less robust effects: older adults had smaller response amplitudes than younger adults, and older adults had greater intersubject variability than younger adults.

Across a variety of behavioral tasks, older adults have worse performance and increased intrasubject variability compared with young adults (Hultsch et al., 2002, West et al., 2002, Murphy et al., 2007, MacDonald et al., 2012, Lovden et al., 2013). Older adults with mild dementia show more intrasubject variability than healthy age-matched controls (MacDonald et al., 2006) and healthy older adults with more trial-to-trial variability showed greater cognitive declines over time (Lovden et al., 2007).

While a link between increased neural (BOLD fMRI) variability and increased behavioral variability is sensible on its face, the precise link between the two is a matter of speculation. With increased neural variability the distribution of responses in a population of neurons to a given stimulus would become wider, which in turn would make it harder for the brain to differentiate between the possible stimuli that evoked the response (Li et al., 2001).

Consistent with this idea, decreased stimulus specificity has been observed in single cell recordings of older cats and non-human primates (Schmolesky et al., 2000, Hua et al., 2006, Liang et al., 2010). Older adults are particularly impaired in perceiving speech if it is embedded in artificially generated noise (Dubno et al., 1984, Humes, 1996, Sommers et al., 2005, Gosselin & Gagne, 2011). The increased noise in the stimulus could exacerbate the effects of increased neural variability, which could be considered “neural noise”. We did not test older subjects using noisy audiovisual speech, the type of speech on which they are most impaired. Therefore, we could not directly compare the increased neural variability we observed in older subjects with the decreased performance of recognizing speech in noise that they are known to have. Future studies using noisy stimuli would be expected to produce poorer performance and reveal differences between subjects correlated with BOLD variability.

Older adult subjects with poor auditory-only or audiovisual syllable identification performance were excluded from the study. This resulted in a group of high-functioning older adults with very nearly identical behavioral scores to the younger adults, which still revealed a difference in neural variability across age groups. Just as high behavioral variability in high-functioning older adults predicts later declines in cognitive function (Lovden et al., 2007), increased neural variability might also provide a similar marker for future cognitive deficits in the absence of any behavioral differences.

Neural variability at early stages of cortical sensory processing might be compounded by neural variability at decision layers higher in the cortical

hierarchy. Speech perception involves categorical judgments about the identity of each syllable. Neuronal variability could impair these decisions, an effect that may be even more important than added sensory noise (Beck et al., 2012).

Neuronal variability may also differentially affect the ability to make both fine and coarse discriminations. Low levels of neuronal variability favor fine discriminations performed at locations in stimulus space in which neuronal selectivity changes rapidly, while high levels of neuronal variability favor coarse discriminations performed at locations in stimulus space where neuronal responses are maximal (Butts & Goldman, 2006). Therefore, increased neuronal variability with aging might impair fine discrimination while leaving coarse discriminations relatively intact.

While our literature search did not reveal any fMRI studies examining multisensory speech perception in healthy older adults, there have been a number of fMRI studies comparing young and older adults in other tasks. The preponderance of studies have reported more variability in older adults, as we observed in our dataset. D'Esposito et al. (1999) measured BOLD responses in motor cortex to a bilateral button press cued by the appearance of a briefly presented white circle. Greater intrasubject variability in older adults compared to younger adults was observed; there were no differences in response amplitude and only a small increase in intersubject variability. Huettel et al. (2001) measured responses in visual cortex to checkerboard stimuli with no behavioral task. Greater intrasubject and intersubject variability in older adults was observed, with no difference in amplitude of response. Samanez-Larkin et al.

(2010) found that healthy older adults exhibited suboptimal decision-making on a financial investment task compared to younger adults and also exhibited greater temporal variability in the nucleus accumbens.

In contrast to these studies that reported more variability in older adults, Garrett et al. (2012) reported *less* neural variability in older adults using a variety of complex cognitive tasks. One possible explanation for this result is that the analysis of Garrett et al. used a measure derived from multivariate voxel pattern analysis (MVPA) to measure variability across all brain voxels. MVPA analyses and traditional univariate analyses such as ours may give conflicting or even contradictory results. The explanation for these discrepancies is a matter of debate (Jimura & Poldrack, 2012).

A potential confound in BOLD fMRI studies of older populations is vascular changes with age (Fang, 1976). However, studies that directly measure neuronal activity also find age-related changes in variability. Anderson et al. (2012) presented auditory syllables to healthy older adults and measured the auditory brain stem response, an electrophysiological measure of neuronal activity that is not influenced by the vasculature, and found greater variability in older adults. In a study of non-human primates, single neuron responses in V1 and MT of older monkeys had greater variability than in younger monkeys (Yang et al., 2009). These findings suggest that the variability differences in our study have a neuronal component in addition to any possible vascular sources.

Although the results presented here concur with previous results in the literature, including neural processing in older adults as well as

electrophysiological results from older experimental animals, it will be important to definitively demonstrate that variability in the BOLD signal has a distinctively neural origin and is not the result of changes to neural-vascular coupling in aging. To do this, physiological measurements (*e.g.* heart rate and respiration) can be recorded during the fMRI experiment for both younger and older subjects and RETROICOR can be used to filter respiration and cardiac induced noise (Glover et al., 2000). If older adults still have greater BOLD signal variability after corrections for differences in heart rate and respiration, this would strengthen the claim that variability in the BOLD signal is neural as opposed to vascular.

Intergroup differences in fMRI studies may also be driven by differences in head movements. This is especially of concern in studies of resting state functional connectivity (Power et al., 2012, Van Dijk et al., 2012). In a resting state study, movements can introduce correlations in the MR time series between distant brain areas that can be wrongly interpreted as evidence for functional connectivity. However, in our study, we did not perform a functional connectivity analysis, nor did we analyze resting state data. In a task-based study such as ours, averaging the response to multiple stimuli can reduce movement effects, since head movements and stimulus presentation are independent. We used five different methods (including the motion scrubbing procedure suggested by Power and colleagues) to account for differences in head motion, and found no effect on our main results. Consistent with these analyses, two previous studies (Huettel et al. (2001) and D'Esposito et al. (1999)) did not find a correlation between head motion and BOLD signal variability.

Our secondary findings of lower response amplitude and greater intersubject variability in older adults were more pronounced in the whole-brain than the ROI analyses. One possibility is that the anatomical templates used for normalization (only necessary for the whole-brain analysis) can introduce group differences (Samanez-Larkin & D'Esposito, 2008). For instance, if older brains atrophy or undergo other morphological changes that vary from individual to individual, they will be less likely to align with the template. Then, the same functional brain region will lie in different locations in standard space in different subjects, leading to increased intersubject variability and decreased response amplitude that has an anatomical, rather than a functional origin. The ROI analysis is less susceptible to this problem because the regions are defined functionally, not based on their location in standard space.

Our most robust finding was of greater intrasubject variability in older adults. This finding was true in both the ROI and whole-brain analysis. Among the physical changes to the aging brain, a decrease in myelination has been observed (Lu et al., 2011, Kerchner et al., 2012). These decreases in white matter integrity could lead to increases in neuronal variability by preventing neurons from firing consistently even with the same sensory input. Better understanding the neural sources of this variability and its behavioral consequences may help in designing strategies to ameliorate declines in speech perception, one of our most important cognitive functions.

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

Speech perception is a critical cognitive function, but the brain mechanisms that support this process are not entirely understood, especially after changes to the brain after stroke or during healthy aging. In Chapter 2, I presented the results of a case study of patient SJ, which illustrated how the brain can support multisensory integration following damage to the left STS. Case studies can be particularly useful when a patient presents with a unique but interpretable deficit that can illuminate something about the organization of the brain (Editorial, 2004).

Patient SJ suffered a cerebrovascular insult damaged the left STS and surrounding tempo-parietal area. Although the left STS seems to be a critical hub in the multisensory speech perception network in healthy adults (Scott & Johnsrude, 2003, Beauchamp, 2005, Miller & D'Esposito, 2005, Stevenson & James, 2009, Beauchamp et al., 2010, Nath & Beauchamp, 2011, 2012), SJ demonstrated multisensory integration abilities through behavioral testing at a level similar to healthy controls. My fMRI study provided evidence that SJ's right STS now supports multisensory integration in speech, as demonstrated by greater multisensory cortex and increased response amplitude in the right STS compared to healthy controls.

This is remarkable because recovered function after a stroke is a desired, but not always actual outcome. Another stroke patient, AWF, acquired a selective impairment in audiovisual processing as the result of an unknown neurologic event (Hamilton et al., 2006). An MRI did not reveal any noticeable damage but a single positron emission computed tomography (SPECT) study

revealed hypoperfusion in parietal cortex bilaterally. AWG performed better on a number of tasks during the unimodal (auditory information only) but not bimodal (auditory and visual information available) conditions, and he did not perceive the McGurk effect. Although this case study did not include functional MRI testing, it would be interesting to contrast the neural response to audiovisual speech in AWG to SJ. It is possible that due to the diffuse nature of the stroke damage bilaterally, AWG's brain was unable to compensate for the damaged area with increased function in healthy cortex.

Like the left STS, the right STS is also anatomically connected to auditory cortex and visual cortex, but may play a less important role in audiovisual integration in speech in healthy subjects. After brain damage, these redundant connections may be utilized to recover lost functions of the damaged cortex. One area that is currently under investigated is the role of the right STS in healthy subjects. Although approximately 93% of healthy adults right-handed adults (Knecht et al., 2000) and 78% of left-handed adults (Szaflarski et al., 2002) show left hemisphere dominance in language perception, the role of the right hemisphere in language processing in healthy adults is not entirely known. Healthy subjects who show less language lateralization (more bilateral processing) are less susceptible to unilateral virtual lesions with TMS (Knecht et al., 2002). Because other studies have shown that disrupting the right hemisphere of recovered aphasic patients results in language impairments (Kinsbourne, 1971, Czopf, 1979, Winhuisen et al., 2005, Turkeltaub et al., 2011),

I predict that disruption of the right STS with TMS would have a much greater impact on SJ's remaining multisensory abilities than healthy controls.

These individual differences in where and how language is processed in the brain may provide some patients with a greater predisposition for recovery than others. In the TMS study by Beauchamp et al. (2010), only left STS activity was disrupted. A follow-up study should also include the multisensory cortex in right STS as an additional TMS site. I predict that the impact of TMS on the right STS of healthy controls would be less than the disruption of left STS activity, but still might decrease McGurk perception. Furthermore, the impact of TMS on McGurk perception in the right hemisphere in individual subjects might be related to degree of language lateralization in each subject.

Future work should track stroke patients through recovery, conducting fMRI and behavioral testing at multiple time points. For those patients that eventually show strong recovery after a stroke, I would predict that the area of multisensory cortex should increase commensurate with improvement on standardized language testing. In patients with Alzheimer's disease, right hemisphere recruitment of language processing correlated positively with performance accuracy in a naming test (Nelissen et al., 2011), demonstrating that certain neural measures may provide an additional marker for disease progression or recovery. In the future, these neural measures could be used in conjunction with traditional neuropsychological testing. Although Beauchamp et al. (2010) used TMS to produce inhibitory effects, altering the frequency and intensity of TMS pulses can also produce excitatory effects (Pascual-Leone et

al., 1994). Techniques that enhance right hemisphere recruitment, like repetitive TMS (rTMS) or transcranial direct current stimulation (tDCS), may be useful to augment language recovery in patients with left hemisphere damage (Medina et al., 2012).

In Chapter 3, I discussed neural responses to audiovisual speech in healthy older and younger adults. I sought to understand some of the neural mechanisms underlying multisensory speech perception in healthy aging. In behavioral tasks, older adults show much greater variability in their trial-to-trial performance. Measures of behavioral variability in healthy aging have shown that this greater variability in performance is associated with greater declines in cognition. For example, intraindividual variability on a simple reaction time task in healthy older adults was shown to be a strong predictor of information retained after a 1 week interval: those with greater intraindividual variability forgot more information than those with less performance variability (Papenberg et al., 2011). To determine if intraindividual variability on neuropsychological testing measures were associated with performance deficits on a more real-world situation, Kennedy et al. (2013) examined performance in a flight simulator in healthy middle-aged and older pilots and found that high intraindividual variability on neuropsychological tests was associated with worse performance on communication, emergency detection, and traffic avoidance while in the flight simulation.

The results of the work presented here show that there is a neural counterpart to this behavioral observation: in older adults there is greater

intraindividual neural variability compared to younger adults. This was true both in specific nodes in the multisensory speech perception network (auditory cortex, visual cortex, and the STS) and across areas that were significantly active during audiovisual speech perception. This, in addition to evidence of increased neural variability from other sensory tasks (D'Esposito et al., 1999, Huettel et al., 2001, Samanez-Larkin et al., 2010), suggests that increased neural variability may be a general feature of the aging brain.

The fact that increased behavioral variability is reliably correlated with declines in cognitive function suggests that there is a link between behavioral variability and cognition. The ability for the brain to receive, filter, and process sensory information is unlikely to be independent from cognitive processes, which also rely on the processing and transforming of information. If this is the case, then the study of variability is important both in its own right and as a window into cognition. The relationship between behavioral variability and neural variability is currently unknown. Preliminary work has suggested two possible mechanisms, including decreased binding of the dopamine D2 receptor (MacDonald et al., 2009) Specifically, decreased D2 receptor binding in the orbitofrontal cortex, anterior cingulate cortex, and hippocampus is correlated with increases in intraindividual standard deviation on cognitive tasks in middle-aged adults. Additionally, decreased white matter integrity (Lu et al., 2011, Kerchner et al., 2012) is correlated with decreases in information processing speed in the aging brain. Decreased myelination and receptor binding may

reduce the efficiency of the propagation of the neural signal, and thus may play a role in the neural and behavioral changes found in the aging brain.

In future studies of healthy aging, it will be important to combine behavioral measures (such as a hearing speech-in-noise task) with fMRI. The Quick Speech-in-Noise (SIN) test (Killion et al., 2004) tests language comprehension at multiple signal-to-noise ratios (SNRs), and it would be interesting to compare speech-in-noise performance with intrasubject BOLD variability. If behavioral variability and neural variability are related, then behavioral measures should correlate with our neural measure of variability, such that older adults with greater neural variability should also show worse performance on sensory behavioral tasks like Quick SIN. This variability may also be correlated with stimulus noise (greater neural variability with decreasing SNR), reflecting both the internal neural noise and external stimulus noise. If we are better able to understand the neural mechanisms of this behavioral variability and its potentially detrimental effects on behavioral performance, this may help in designing strategies to increase speech perception (and other kinds of sensory processing) in older adults.

BIBLIOGRAPHY

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N (2012) Aging affects neural precision of speech encoding. The Journal of Neuroscience 32:14156-14164.

Argall BD, Saad ZS, Beauchamp MS (2006) Simplified intersubject averaging on the cortical surface using SUMA. Human Brain Mapping 27:14-27.

Baum SH, Martin RC, Hamilton AC, Beauchamp MS (2012) Multisensory speech perception without the left superior temporal sulcus. NeuroImage 62:1825-1832.

Beauchamp MS (2005) See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Current Opinion in Neurobiology 15:145-153.

Beauchamp MS, Lee KE, Argall BD, Martin A (2004) Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41:809-823.

Beauchamp MS, Nath AR, Pasalar S (2010) fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. Journal of Neuroscience 30:2414-2417.

Beauchamp MS, Yasar NE, Kishan N, Ro T (2007) Human MST but not MT responds to tactile stimulation. Journal of Neuroscience 27:8261-8267.

Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A (2012) Not noisy, just wrong: the role of suboptimal inference in behavioral variability. Neuron 74:30-39.

Bielak AA, Cherbuin N, Bunce D, Anstey KJ (2013) Intraindividual Variability Is a Fundamental Phenomenon of Aging: Evidence From an 8-Year Longitudinal Study Across Young, Middle, and Older Adulthood. Dev Psychol.

Blasi V, Young AC, Tansy AP, Petersen SE, Snyder AZ, Corbetta M (2002) Word retrieval learning modulates right frontal cortex in patients with left frontal damage. Neuron 36:159-170.

Blau V, van Atteveldt N, Formisano E, Goebel R, Blomert L (2008) Task-irrelevant visual letters interact with the processing of speech sounds in heteromodal and unimodal cortex. Eur J Neurosci 28:500-509.

Braitenberg V, Schuez A (1998) Cortex: Statistics and geometry of neuronal connectivity. Berlin: Springer.

Brant LJ, Fozard JL (1990) Age changes in pure-tone hearing thresholds in a longitudinal study of normal human aging. J Acoust Soc Am 88:813-820.

Buckner RL, Corbetta M, Schatz J, Raichle ME, Petersen SE (1996) Preserved Speech Abilities and Compensation Following Prefrontal Damage. Proceedings Of the National Academy Of Sciences Of the United States Of America 93:1249-1253.

Butts DA, Goldman MS (2006) Tuning curves, neuronal variability, and sensory coding. PLoS Biol 4:e92.

Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol 10:649-657.

Campbell R, Garwood J, Franklin S, Howard D, Landis T, Regard M (1990) Neuropsychological studies of auditory-visual fusion illusions. Four case studies and their implications. Neuropsychologia 28:787-802.

Cao Y, Vikingstad EM, George KP, Johnson AF, Welch KM (1999) Cortical language activation in stroke patients recovering from aphasia with functional MRI. Stroke 30:2331-2340.

Carreiras M, Seghier ML, Baquero S, Estevez A, Lozano A, Devlin JT, Price CJ (2009) An anatomical signature for literacy. Nature 461:983-986.

Cerella J, Hale S (1994) The rise and fall in information-processing rates over the life span. Acta Psychol (Amst) 86:109-197.

Champoux F, Tremblay C, Mercier C, Lassonde M, Lepore F, Gagne JP, Theoret H (2006) A role for the inferior colliculus in multisensory speech integration. Neuroreport 17:1607-1610.

Colin C, Radeau M, Deltenre P (2005) Top-down and bottom-up modulation of audiovisual integration in speech. European Journal of Cognitive Psychology 17:541-560.

Cox RW (1996) AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research 29:162-173.

Crawford JR, Howell, D.C. (1998) Comparing an Individual's Test Scores Against Norms Derived from Small Samples. The Clinical Neuropsychologist 12:482-486.

Czopf D (1979) The role of the non-dominant hemisphere in speech recovery. Aphasia Apraxia Agnosia 2:27-33.

D'Esposito M, Zarahn E, Aguirre GK, Rypma B (1999) The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. NeuroImage 10:6-14.

Dale AM (1999) Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109-114.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage 9:179-194.

Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage 53:1-15.

Divenyi PL, Stark PB, Haupt KM (2005) Decline of speech understanding and auditory thresholds in the elderly. J Acoust Soc Am 118:1089-1100.

Dronkers NF, Wilkins DP, Van Valin RD, Jr., Redfern BB, Jaeger JJ (2004) Lesion analysis of the brain areas involved in language comprehension. Cognition 92:145-177.

Dubno JR, Dirks DD, Morgan DE (1984) Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am 76:87-96.

Editorial (2004) When once is enough. Nature Neuroscience 7:93.

Fang HCH (1976) Observations on aging characteristics of cerebral blood vessels, macroscopic and microscopic features. In: Neurobiol

Aging(Terry, R. D. and Gershon, S., eds), pp 155-166 New York: Raven Press.

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. NeuroImage 9:195-207.

Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004) Automatically parcellating the human cerebral cortex. Cerebral Cortex 14:11-22.

Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189-198.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med 33:636-647.

Fox PT, Raichle ME (1986) Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. Proceedings Of the National Academy Of Sciences Of the United States Of America 83:1140-1144.

Fox PT, Raichle ME, Mintun MA, Dence C (1988) Nonoxidative glucose consumption during focal physiologic neural activity. Science 241:462-464.

Frassinetti F, Bolognini N, Bottari D, Bonora A, Ladavas E (2005) Audiovisual integration in patients with visual deficit. J Cogn Neurosci 17:1442-1452.

Friston KJ (2009) Modalities, modes, and models in functional neuroimaging. Science 326:399-403.

Friston KJ, Rotshtein P, Geng JJ, Sterzer P, Henson RN (2006) A critique of functional localisers. NeuroImage 30:1077-1087.

Froyen D, van Atteveldt N, Blomert L (2010) Exploring the Role of Low Level Visual Processing in Letter-Speech Sound Integration: A Visual MMN Study. Front Integr Neurosci 4:9.

Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2012) The Modulation of BOLD Variability between Cognitive States Varies by Age and Processing Speed. Cerebral Cortex.

Glover GH, Li TQ, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn Reson Med 44:162-167.

Gordon GR, Mulligan SJ, MacVicar BA (2007) Astrocyte control of the cerebrovasculature. Glia 55:1214-1221.

Gosselin PA, Gagne JP (2011) Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. Int J Audiol 50:786-792.

Grant KW, Seitz PF (2000) The use of visible speech cues for improving auditory detection of spoken sentences. Journal of the Acoustical Society of America 108:1197-1208.

Hamilton RH, Shenton JT, Coslett HB (2006) An acquired deficit of audiovisual speech processing. Brain Lang 98:66-73.

Hedden T, Gabrieli JD (2004) Insights into the ageing mind: a view from cognitive neuroscience. Nat Rev Neurosci 5:87-96.

Hickok G, Costanzo M, Capasso R, Miceli G (2011) The role of Broca's area in speech perception: evidence from aphasia revisited. Brain Lang 119:214-220.

Hua T, Li X, He L, Zhou Y, Wang Y, Leventhal AG (2006) Functional degradation of visual cortical cells in old cats. Neurobiol Aging 27:155-162.

Huettel SA, Singerman JD, McCarthy G (2001) The effects of aging upon the hemodynamic response measured by functional MRI. NeuroImage 13:161-175.

Huettel SA, Song AW, McCarthy G (2009) Functional Magnetic Resonance Imaging. Sunderland, MA: Sinauer Associates.

Hultsch DF, MacDonald SW, Dixon RA (2002) Variability in reaction time performance of younger and older adults. J Gerontol B Psychol Sci Soc Sci 57:P101-115.

Humes LE (1996) Speech understanding in the elderly. J Am Acad Audiol 7:161-167.

Jenkins L, Myerson J, Joerding JA, Hale S (2000) Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. Psychol Aging 15:157-175.

Jimura K, Poldrack RA (2012) Analyses of regional-average activation and multivoxel pattern information tell complementary stories. Neuropsychologia 50:544-552.

Kennedy Q, Taylor J, Heraldez D, Noda A, Lazzeroni LC, Yesavage J (2013) Intraindividual variability in basic reaction time predicts middle-aged and older pilots' flight simulator performance. J Gerontol B Psychol Sci Soc Sci 68:487-494.

Kerchner GA, Racine CA, Hale S, Wilheim R, Laluz V, Miller BL, Kramer JH (2012) Cognitive processing speed in older adults: relationship with white matter integrity. PLoS One 7:e50425.

Kherif F, Josse G, Seghier ML, Price CJ (2009) The main sources of intersubject variability in neuronal activation for reading aloud. J Cogn Neurosci 21:654-668.

Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S (2004) Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am 116:2395-2405.

Kinsbourne M (1971) The minor cerebral hemisphere as a source of aphasic speech. Arch Neurol 25:302-306.

Knecht S, Deppe M, Drager B, Bobe L, Lohmann H, Ringelstein E, Henningsen H (2000) Language lateralization in healthy right-handers. Brain 123 (Pt 1):74-81.

Knecht S, Floel A, Drager B, Breitenstein C, Sommer J, Henningsen H, Ringelstein EB, Pascual-Leone A (2002) Degree of language lateralization determines susceptibility to unilateral brain lesions. Nature Neuroscience 5:695-699.

Kriegeskorte N, Simmon WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nature Neuroscience 12:535-540.

Lewis JW, Van Essen DC (2000) Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. Journal of Comparative Neurology 428:112-137.

Li SC, Lindenberger U, Sikstrom S (2001) Aging cognition: from neuromodulation to representation. Trends Cogn Sci 5:479-486.

Liang Z, Yang Y, Li G, Zhang J, Wang Y, Zhou Y, Leventhal AG (2010) Aging affects the direction selectivity of MT cells in rhesus monkeys. Neurobiol Aging 31:863-873.

Logothetis NK, Wandell BA (2004) Interpreting the BOLD signal. Annu Rev Physiol 66:735-769.

Lovden M, Li SC, Shing YL, Lindenberger U (2007) Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: longitudinal data from the Berlin Aging Study. Neuropsychologia 45:2827-2838.

Lovden M, Schmiedek F, Kennedy KM, Rodrigue KM, Lindenberger U, Raz N (2013) Does variability in cognitive performance correlate with frontal brain volume? NeuroImage 64:209-215.

Lu PH, Lee GJ, Raven EP, Tingus K, Khoo T, Thompson PM, Bartzokis G (2011) Age-related slowing in cognitive processing speed is associated with myelin integrity in a very healthy elderly sample. J Clin Exp Neuropsychol 33:1059-1068.

MacDonald SW, Cervenka S, Farde L, Nyberg L, Backman L (2009) Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. Neuropsychologia 47:2299-2304.

MacDonald SW, Karlsson S, Rieckmann A, Nyberg L, Backman L (2012) Aging-related increases in behavioral variability: relations to losses of dopamine D1 receptors. J Neurosci 32:8186-8191.

MacDonald SW, Nyberg L, Backman L (2006) Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. Trends in Neurosciences 29:474-480.

MacLeod A, Summerfield Q (1990) A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. Br J Audiol 24:29-43.

Massaro D (1998) Data Archive of 5*5+5+5 Expanded Factorial Visual-Auditory Recognition Experiments.

Massaro D, Cohen MM, Gesi A, Heredia R, Tsuzaki M (1993) Bimodal speech perception: an examination across languages. Journal of Phonetics 21:445-478.

Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Woods R, Paus T, Simpson G, Pike B, Holmes C, Collins L, Thompson P, MacDonald D, Iacoboni M, Schormann T, Amunts K, Palomero-Gallagher N, Geyer S, Parsons L, Narr K, Kabani N, Le Goualher G, Boomsma D, Cannon T, Kawashima R, Mazoyer B (2001) A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philosophical Transactions of the Royal Society B: Biological Sciences 356:1293-1322.

McGettigan C, Faulkner A, Altarelli I, Obleser J, Baverstock H, Scott SK (2012) Speech comprehension aided by multiple modalities: Behavioural and neural interactions. Neuropsychologia.

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264:746-748.

Medina J, Norise C, Faseyitan O, Coslett HB, Turkeltaub PE, Hamilton RH (2012) Finding the Right Words: Transcranial Magnetic Stimulation Improves Discourse Productivity in Non-fluent Aphasia After Stroke. Aphasiology 26:1153-1168.

Miller LM, D'Esposito M (2005) Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. Journal of Neuroscience 25:5884-5893.

Mumford JA, Turner BO, Ashby FG, Poldrack RA (2012) Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. NeuroImage 59:2636-2643.

Murphy KJ, West R, Armilio ML, Craik FI, Stuss DT (2007) Word-list-learning performance in younger and older adults: intra-individual performance variability and false memory. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 14:70-94.

Nath AR, Beauchamp MS (2011) Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech. Journal of Neuroscience 31:1704-1714.

Nath AR, Beauchamp MS (2012) A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. NeuroImage 59:781-787.

Nath AR, Fava EE, Beauchamp MS (2011) Neural correlates of interindividual differences in children's audiovisual speech perception. Journal of Neuroscience 31:13963-13971.

Nelissen N, Dupont P, Vandenbulcke M, Tousseyn T, Peeters R, Vandenberghe R (2011) Right hemisphere recruitment during language processing in frontotemporal lobar degeneration and Alzheimer's disease. J Mol Neurosci 45:637-647.

Nizar K, Uhlirova H, Tian P, Saisan PA, Cheng Q, Reznichenko L, Weldy KL, Steed TC, Sridhar VB, MacDonald CL, Cui J, Gratiy SL, Sakadzic S, Boas DA, Beka TI, Einevoll GT, Chen J, Masliah E, Dale AM, Silva GA,

Devor A (2013) In vivo stimulus-induced vasodilation occurs without IP3 receptor activation and may precede astrocytic calcium increase. J Neurosci 33:8411-8422.

Ogawa S, Lee TM (1990) Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation. Magn Reson Med 16:9-18.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97-113.

Papenberg G, Backman L, Chicherio C, Nagel IE, Heekeren HR, Lindenberger U, Li SC (2011) Higher intraindividual variability is associated with more forgetting and dedifferentiated memory functions in old age. Neuropsychologia 49:1879-1888.

Pascual-Leone A, Valls-Sole J, Wassermann EM, Hallett M (1994) Responses to rapid-rate transcranial magnetic stimulation of the human motor cortex. Brain 117 (Pt 4):847-858.

Peich MC, Husain M, Bays PM (2013) Age-related decline of precision and binding in visual working memory. Psychol Aging 28:729-743.

Poldrack RA (2007) Region of interest analysis for fMRI. Soc Cogn Affect Neurosci 2:67-70.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59:2142-2154.

Price LL (1963) Threshold testing with Bekesy audiometer. J Speech Hear Res 6:64-69.

Raij T, Uutela K, Hari R (2000) Audiovisual integration of letters in the human brain. Neuron 28:617-625.

Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007) Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cerebral Cortex 17:1147-1153.

Saad ZS, Glen DR, Chen G, Beauchamp MS, Desai R, Cox RW (2009) A new method for improving functional-to-structural MRI alignment using local Pearson correlation. NeuroImage 44:839-848.

Samanez-Larkin GR, D'Esposito M (2008) Group comparisons: imaging the aging brain. Soc Cogn Affect Neurosci 3:290-297.

Samanez-Larkin GR, Kuhnen CM, Yoo DJ, Knutson B (2010) Variability in nucleus accumbens activity mediates age-related suboptimal financial risk taking. J Neurosci 30:1426-1434.

Saxe R, Brett M, Kanwisher N (2006) Divide and conquer: a defense of functional localizers. NeuroImage 30:1088-1096; discussion 1097-1089.

Schmid G, Thielmann A, Ziegler W (2009) The influence of visual and auditory information on the perception of speech and non-speech oral movements in patients with left hemisphere lesions. Clin Linguist Phon 23:208-221.

Schmolesky MT, Wang Y, Pu M, Leventhal AG (2000) Degradation of stimulus selectivity of visual cortical cells in senescent rhesus monkeys. Nature Neuroscience 3:384-390.

Scott SK, Johnsrude IS (2003) The neuroanatomical and functional organization of speech perception. Trends in Neurosciences 26:100-107.

Seltzer B, Cola MG, Gutierrez C, Massee M, Weldon C, Cusick CG (1996) Overlapping and nonoverlapping cortical projections to cortex of the superior temporal sulcus in the rhesus monkey: double anterograde tracer studies. Journal of Comparative Neurology 370:173-190.

Skudlarski P, Constable RT, Gore JC (1999) ROC analysis of statistical methods used in functional MRI: individual subjects. NeuroImage 9:311-329.

Sommers MS, Tye-Murray N, Spehar B (2005) Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. Ear Hear 26:263-275.

Stein BE, Meredith MA (1993) The Merging of the Senses: MIT Press.

Stevenson RA, James TW (2009) Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. NeuroImage 44:1210-1223.

Stevenson RA, VanDerKlok RM, Pisoni DB, James TW (2011) Discrete neural substrates underlie complementary audiovisual speech integration processes. NeuroImage 55:1339-1345.

Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. Journal of the Acoustic Society of America 26:212-215.

Szaflarski JP, Binder JR, Possing ET, McKiernan KA, Ward BD, Hammeke TA (2002) Language lateralization in left-handed and ambidextrous people: fMRI data. Neurology 59:238-244.

Thomas C, Altenmueller, E., Marckmann, G., Kahrs, J., Dichgans, J. (1997) Language processing in aphasia: changes in lateralization patterns during recovery reflect cerebral plasticity in adults. Electroencephalography and Clinical Neurophysiology 102:86-97.

Turkeltaub PE, Coslett HB, Thomas AL, Faseyitan O, Benson J, Norise C, Hamilton RH (2011) The right hemisphere is not unitary in its role in aphasia recovery. Cortex.

van Atteveldt N, Formisano E, Goebel R, Blomert L (2004) Integration of letters and speech sounds in the human brain. Neuron 43:271-282.

van Atteveldt NM, Blau VC, Blomert L, Goebel R (2010) fMR-adaptation indicates selectivity to audiovisual content congruency in distributed clusters in human superior temporal cortex. BMC Neurosci 11:11.

Van Dijk KR, Sabuncu MR, Buckner RL (2012) The influence of head motion on intrinsic functional connectivity MRI. NeuroImage 59:431-438.

Vandermorris S, Murphy KJ, Troyer AK (2013) Age-related elevations in intraindividual variability on associative memory tasks. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 20:722-734.

Vul E, Harris C, Winkielman P, Pashler H (2009) Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Congition. Perspectives on Psychological Science 4:274-290.

West R, Murphy KJ, Armilio ML, Craik FI, Stuss DT (2002) Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. Brain Cogn 49:402-419.

Wilson M (1988) MRC psycholinguistic database: Machine-usable dictionary, version 2.00. Behavioral Research Methods, Instruments, & Computers 20:6-10.

Winhuisen L, Thiel A, Schumacher B, Kessler J, Rudolf J, Haupt WF, Heiss WD (2005) Role of the contralateral inferior frontal gyrus in recovery of language function in poststroke aphasia: a combined repetitive transcranial magnetic stimulation and positron emission tomography study. Stroke 36:1759-1763.

Wylie GR, Genova H, Deluca J, Chiaravalloti N, Sumowski JF (2012) Functional magnetic resonance imaging movers and shakers: Does subject-movement cause sampling bias? Human Brain Mapping.

Xiong J, Gao J, Lancaster J, Fox P (1995) Clustered pixels analysis for functional MRI activation studies of the human brain. Human Brain Mapping 3:287-301.

Yang Y, Liang Z, Li G, Wang Y, Zhou Y (2009) Aging affects response variability of V1 and MT neurons in rhesus monkeys. Brain Res 1274:21-27.

Youse KM, Cienkowski KM, Coelho CA (2004) Auditory-visual speech perception in an adult with aphasia. Brain Injury 18:825-834.

VITA

Sarah Haller Baum was born in Williamsburg, Virginia on March 9, 1987, the daughter of Kimberly Jean Haller Baum and Robert Stephen Baum. After completing her work at South Lakes High School in Reston, Virginia in 2005, she entered the College of William & Mary in Williamsburg, Virginia. She received the degree of Bachelor of Science with a major in neuroscience from William & Mary in May 2009. In August of 2009, she entered the Ph.D. Program in Neuroscience at The University of Texas Health Science Center at Houston Graduate School of Biomedical Sciences.

Permanent address:
2427 Freetown Drive
Reston, VA 20191