ADAPTIVE PROCESSES IN SPEECH PERCEPTION: CONTRIBUTIONS FROM CEREBRAL AND CEREBELLAR CORTICES

by

Sara Guediche

BS/MS Chemistry, Bucknell University, 1999

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

School of Arts and Sciences

This dissertation was presented

by

Sara Guediche

It was defended on

October 27, 2009

and approved by

Marc Sommer, Associate Professor, Neuroscience

Peter L. Strick, Professor, Neurobiology, Psychiatry

Mark E. Wheeler, Assistant Professor, Neuroscience and Psychology

Lori L. Holt, Associate Professor, Psychology, Carnegie Mellon University

Steven L. Small, Professor, Neurology, Psychology, University of Chicago

Dissertation Advisor: Julie A. Fiez, Professor Neuroscience and Psychology

ADAPTIVE PROCESSES IN SPEECH PERCEPTION: CONTRIBUTIONS FROM CEREBRAL AND CEREBELLAR CORTICES

Sara Guediche, PhD

University of Pittsburgh, 2009

In the sensorimotor domain, adaptation to distorted sensory input has been well-characterized and is largely attributed to learning mechanisms in the cerebellum that adjust motor output to achieve the same desired sensory outcome. Our interest in the role of the cerebellum in cognitive processes has led us to question whether it also contributes to adaptation in tasks that do not require voluntary motor output. Speech perception is a domain where there exist many examples of adaptation that are guided by both sensory and cognitive processes, without intentional motor involvement. Thus, we investigated behavioral and neural characteristics of speech perception adaptation to spectrally distorted words using a sophisticated noise-vocoded speech manipulation that mimics cochlear implants. We demonstrated that adaptation to spectrally distorted words can be achieved without explicit feedback by either gradually increasing the severity of the distortion or by using an intermediate distortion during training. We identified regions in both the cerebellar and cerebral cortex that showed differences in neural responses before and after training. In the cerebellum, this included regions in lobes V and VI, and Crus I. In the cerebrum, this included regions in the inferior frontal gyrus, the superior temporal sulcus, and the posterior inferior/middle temporal gyrus. In some of these regions, we further found changes in the magnitude of the neural responses that corresponded to the degree of behavioral improvements in performance. To gain some insight into the nature of the interactions between cerebral and cerebellar cortices and the types of representations involved in speech perception adaptation, we conducted a simple functional connectivity analysis using cerebellar seed regions of interest. We found interactions between the cerebellum and cerebral cortex that were dependent on the location of the cerebellar region. Overall, our behavioral and functional neuroimaging results point to cerebellar involvement in speech perception adaptation, and we conclude with a discussion of the learning mechanisms and neuroanatomical pathways that may support such plasticity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

There are no words that can describe the gratitude and high level of respect that I have for my advisor, Dr. Julie Fiez. She is a brilliant investigator whose commitment to research, teaching, and mentoring has encouraged me to grow as a scientist and as a person. She tailors her mentoring style to meet the individual needs of each student and somehow, no matter how many students she has or how busy she is, she always finds a way to make time for everyone. My experience in Julie's lab goes far beyond the scope of the work presented in this dissertation. The diverse nature of Julie's research program enabled me to be involved in a number of different projects, which significantly enriched my graduate school experience. In addition, the members of her lab create a pleasant and nurturing environment that breeds success. I am especially thankful to my first officemate, Dr. Elizabeth Tricomi, who provided me with a great deal of intellectual guidance during my first few years and Dr. Gal Ben-Yehudah, a post-doc in our lab who was always willing to discuss ideas. I would also like to thank Michelle Moore, my current officemate and Corrine Durisko, our lab coordinator for their daily support.

I was fortunate to have not just one, but two amazing mentors during my graduate experience. Dr. Lori Holt provided me a wealth of additional training and resources. The work in Chapter 2 was completed under her guidance. I am very grateful for her invaluable scientific contributions, her time, and most of all her contagious enthusiasm. Julie and Lori are both amazing role models that I sincerely admire.

I am also thankful to all my other committee members, my committee chair Marc Sommer, Dr. Peter Strick, and Dr. Mark Wheeler, for their guidance and intellectual input. I respect them greatly for encouraging me to think more critically and deeply about each issue and for enhancing my graduate school training. I am also grateful to Dr. Steven Small who graciously agreed to be my outside examiner. I admire his work immensely and am thankful for the opportunity to get to know him and discuss my work with him.

Both the CNUP and the CNBC provided me with a unique opportunity to interact with exceptional students and faculty who are invested in the success of their students. The Neuroscience department staff including Joan Blaney and Marlene Nieri made all the other aspects of graduate school easily manageable. One of the most enjoyable facets of graduate student life is the interaction with fellow students. I especially appreciate the constant support of my friends and colleagues who gave me the strength to persevere through many obstacles. I would like to specifically mention Dr. Lesley Colgan and Patryk Laurent, two classmates, friends, and colleagues with whom I was able to share my successes and failures throughout the entire time.

Most importantly, the infinite encouragement and love of my family members helped in ways that I cannot even begin to describe. I dedicate this thesis to my mother Judy Guediche, a devoted mother and teacher who instilled in me the value of education from an early age. Finally, I give a special thank you to my wonderful husband, Amin, who inspires me and brings so much joy into my life.

1.0 INTRODUCTION

Although cerebellar involvement in cognitive processes has been established, the functional relevance of its role in cognitive tasks still remains unclear. Historically, functions attributed to the cerebellum have included motor coordination, timing, and error-driven learning (Albus, 1971; Marr, 1969; Bloedel & Bracha, 1998; Ivry R. B., 2004; Braitenberg, 1967; Thompson & Steinmetz, 2009). Based on the homogeneity of cerebellar cellular organization, investigators have speculated that similar cerebellar mechanisms apply to the domain of language (Thach W. , 2007; Daum & Ackermann, 1995; Strick, Dum, & Fiez, 2009; Doya, 2000; Ito, Control of mental activities by internal models in the cerebellum, 2008) and cognition whereas others have attributed the involvement of the cerebellum in these tasks to a motor planning component (Glickstein, 2007). Our goal is to investigate the role of the cerebellum in an adaptation task that does not require any intentional motor output. Specifically, we examined a speech perception adaptation task to spectrally distorted speech.

1.1 CEREBELLAR INVOLVEMENT IN MOTOR PROCESSES

Much of our knowledge about the function of the cerebellum comes from investigations of the vestibular ocular reflex, eyeblink conditioning, and motor coordination (Ghez & Thach, 2000). Detailed investigations of these tasks have shed light on the correspondence between cellular changes in activity and behavior leading to several theoretical proposals about the algorithms computed in the cerebellum and how they contribute to motor output.

Examination of patients with cerebellar damage has provided essential insight into the types of functions that the cerebellum performs. A range of motor disorders, including ataxia, intention tremor, dysarthria, and dysdiadochokinesi, have been associated with cerebellar damage, and support a central role in aspects of motor control including planning, coordination, and rhythm (Ghez & Thach, 2000). Specifically, these disorders have indicated the adaptive role of cerebellar functions. Whereas patients often have difficulty making fine adjustments to movements, their ability to initiate a movement is often left intact.

Cerebellar interactions with motor cortex have been well-defined with cerebellar efferents projecting to the dentate nucleus, then through the ventral thalamus, and then ending in cerebral cortex (Ramnani, 2006). The cerebellum receives input from cerebral cortex via the pontine nuclei. These pathways can originate and end in the same areas to form closed cerebrocerebellar loop (Kelly & Strick, 2003; Dum & Strick, 2003). Closed cerebro-cerebellar loops are thought to be the underlying “units” supporting parallel interactions between different cerebral regions and the cerebellum. Both the anterior portions of the cerebellar cortex in lobes V and VI and in posterior HVIIb and HVIII form such loops with motor cortex and posses somatotopic maps (Grodd, Hulsmann, Lotze, Wildgruber, & Erb, 2001). This basic architecture appears to underlie other cerebro-cerebellar interactions, especially for regions associated with motor-related functions.

The computations performed by the cerebellum emerge from an intricate and consistent cellular organization that comprises a number of different excitatory and inhibitory cell types including granule cells, Purkinje cells, Golgi cells, and stellate/basket cells (Voogd & Glickstein, 1998). Since Purkinje cells supply the only output from the cerebellum, they convey essential signals for motor control. Purkinje cells receive excitatory inputs from parallel fibers that project from the granule cells, and from climbing fibers coming from the inferior olive. Mossy fiber input from cerebral cortex provides input to the granular cell layer. Simultaneous inputs from climbing and parallel fibers can lead to temporary changes in the excitability of the Purkinje cell, resulting in plasticity. Long-term depression (LTD) results from simultaneous activation of a Purkinje-parallel fiber synapse and a climbing fiber synapse for the same Purkinje cells, which may underlie motor learning and adaptation (Ito, 2006), whereas long-term potentiation (LTP) is induced by parallel-fiber stimulation is thought to reset LTD at the same synapses. Thus LTD and LTP may constitute the basic learning mechanism in the cerebellum.

Early theories of cerebellum computation focused on two types of information: Braitenberg developed an influential model that discussed how the cerebellar circuitry could be used to produce internal timing signals, while Marr & Albus focused on how the circuitry could be used to produce error signals that drive motor learning and adaptive plasticity (Braitenberg, 1967; Marr, 1969; Albus, 1971). More recent theories have built upon this initial work to propose that the cerebellum contributes to the computation of prediction errors, supervised learning, and the construction of internal models, and also to speculate that the basic computational properties of the cerebellum may apply to non-motor domains (Doya, 2000; Wolpert, Miall, & Kawato, 1998; Schultz & Dickinson, 2000). Given the importance of timing and error information in the cerebellum, we consider their impact on tasks investigated in the classic motor literature.

1.1.1 Timing

An important aspect of motor control is the execution of precisely timed and coordinated movements. Eye-blink conditioning has been used as a classic example to demonstrate cerebellar contributions to time-dependent information that is necessary for learning. In eye-blink conditioning, a blink that follows a puff of air can be evoked with a conditioned tone stimulus after repeatedly pairing the tone with the puff of air. The timing between the tone and the puff of air is crucial in eliciting the conditioned response (Mauk, Medina, Nores, & Ohyama, 2000). The conditioned response is thought to result from reduced excitatory climbing fiber input to the Purkinje cells for the puff that occurs within a certain time frame of the tone. The decrease in climbing fiber activity, in turn, reduces activity in the Purkinje cells, and ultimately leads to increases in activity in the dentate nucleus and in subsequent motor output.

Exactly how the timing information is processed in the cerebellum remains unclear, though compelling hypotheses have emerged. Prevailing conceptualizations of timing in the cerebellum include delay lines and oscillatory models (Ivry, 2004). The computational instantiation of timing delays or oscillations are both produced at the granular cell layer by manipulating different cell properties of the granule cell or by proposing differences in conduction velocities (such as those of the parallel fibers) (Yamazaki & Tanaka, 2009). Granule cell activity can also be modulated by interactions with Golgi cells, which could in turn affect the parallel fiber inputs to Purkinje cells.

Though our understanding of the mechanisms representing time information is still in progress, time-dependent information provided by the cerebellum appears to be essential for motor control. Damage to the cerebellum can produce specific impairments in timing that lead to difficulties in performing rhythmic movements (Braitenberg, 1967). Clinical studies have also shown that cerebellar patients with lateral and not medial damage are impaired in the perception of timing in movement, showing deficits in tasks that require perceptual discrimination between two movements that differ in their temporal durations (Ivry, Keele, & Diener, 1988). These clinical results have been substantiated by functional neuroimaging studies that show increases in activity in the right lateral Crus I area of the cerebellum during a temporal duration discrimination task (Mathiak , Hertrich, Grodd, & Ackermann, 2004).

1.1.2 Short-term sensorimotor adaptation

When a movement results in an unexpected outcome we are able to quickly adapt our movement to produce the desired outcome. Information from different sensory modalities can inform us about the error of our movement, by providing a way to compare the expected and actual sensory outcome. A classic example of sensorimotor adaptation can be demonstrated by performing a visually-guided reaching task with prism goggles that alter the visual feedback. The distorted visual information, such as a lateral shift in the visual field, causes errors in reaching towards a target. Within only a few trials, small adjustments in motor output compensate for the distorted sensory input and enable a successful reach towards the intended target.

Visually-guided reaches depend on posterior parietal cortex and the cerebellum (Baizer & Glickstein, 1974; Clower, Hoffman, Votaw, Faber, Woods, & Alexander, 1996). The role of the cerebellum is critical to sensorimotor adaptation during a visually-guided reach as indicated by the fact that cerebellar patients show severe deficits in adapting to prism goggles (Ghez & Thach, 2000). This is in contrast to patients who have damage to parietal cortex. One patient with a bilateral parietal lesion demonstrated normal adaptation to a prism deviation, but showed impairments in the initial pointing movement (Pisella, Michel, Grea, Tilikete, Vighetto, &

Rossetti, 2004). The authors argue for dissociation between strategic and adaptive components involved in prism adaptation, with the cerebellum being more involved in adaptive mechanisms, and the parietal cortex being more involved in strategic mechanisms.

Functional neuroimaging studies (PET and fMRI) have provided additional evidence supporting the essential role of the cerebellum in short-term adaptation. For example, Diedrichsen et al. showed that the activity in the cerebellum is modulated by the detection of both target errors (changes in the location of the target) and execution errors (changes in motor control) (Diedrichsen, Hashambhoy, Rane, & Shadmehr, 2005). Similarly, studies examining short-term adaptation in saccades showed climbing fiber activity corresponded with the degree of displacement of a visual target (Soetedjo, Kojima, & Fuchs, 2008).

One hypothesis for how sensorimotor adaptation is achieved is that prediction error is used to drive learning. Forward models have been proposed as a means of predicting sensory consequences from a motor command (Wolpert, Miall, & Kawato, 1998). The predicted sensory outcome of a planned motor action can then be compared to the actual sensory feedback. Resulting errors between the predicted and actual sensory feedback modify the forward model (Wolpert, Miall, & Kawato, 1998). The result is that the forward model now has different sensory predictions for a given motor command, thus the relationship between a motor action and a sensory consequence has been altered. Inverse models "predict" what the motor command should be and compare it to the actual motor command. This comparison will generate "feed-back" errors, and modify the inverse model (Ito, 2008).

The post-adaptation effect that manifests, after sensorimotor adaptation, provides additional evidence that sensorimotor "remapping" can occur. For example, after removal of prism glasses, reaching towards a desired target produces reach errors in the opposite direction, a finding which indicates that the internal representation of the sensorimotor correspondence has been shifted. Wolpert et al. discuss the existing evidence supporting the role of the cerebellum in representing forward and inverse models; they also posit that for skilled motor actions, the cerebellum may represent paired forward and inverse models (Wolpert, Miall, & Kawato, 1998).

1.2 CEREBELLAR INVOLVEMENT IN NON-MOTOR PROCESSES

Until recently, the role of the cerebellum was thought to be limited to motor-related functions. This notion has been reconsidered since the late 1980's, when Leiner and colleagues, noted that the evolutionary development of the lateral hemispheres of the cerebellum paralleled a similar evolution of the cerebral cortex prefrontal areas (Leiner, Leiner, & Dow, 1987). Since these regions are associated with cognitive functions, they proposed that the cerebellum might also be involved in cognitive tasks and suggested that the cerebellum may be performing a more general function that extends beyond the motor domain.

New advances in neuroanatomical research that take advantage of viral tracing techniques have begun to reveal the neuroanatomical circuitry that might underlie cerebellar contributions to cognition. Viral tracers cross multiple synapses, and thus enable specific mapping of the pathways that connect the cerebellar and cerebral cortices. Using these techniques, direct anatomical connections were found between lateral portions of the cerebellar hemispheres and prefrontal cortex. The basic concept of closed loop interactions between the cerebellum and motor cortex also applies to connections between the cerebellum and other non-motor cerebral cortical areas. Furthermore, these pathways provide anatomical evidence that supports cerebellar involvement in non-motor functions (Kelly & Strick, 2003; Dum & Strick, 2003).

The cerebro-cerebellar loop hypothesis suggests that the contributions of a given cerebellar region to any task may be similar, but may depend on its anatomical connectivity to specific cerebral cortical areas involved in the task. In addition to the direct anatomical evidence for connections to prefrontal cortex, functional connectivity studies have shown similar results. Using resting state fMRI, Allen et al. demonstrated functional coherence between the dentate nucleus and prefrontal cortex as well as dentate and parietal cortex (Allen, McColl, Barnard, Ringe, Fleckenstein, & Cullum, 2005). In more recent studies using functional connectivity analyses, specific regions within cerebral cortex were used as seed region to identify connections to the cerebellum; specifically, a study by Kriennan et al used several seed regions in different parts of prefrontal cortex and a seed in motor cortex (Krienen & Buckner, 2009). Kriennan et al. examined preferential correlations between the predefined cerebral cortical regions and the distinct subregions that were found in the cerebellum. Their analyses revealed significant differences in the correlations and that provided additional support for the existence of closed cerebro-cerebellar loops (Krienen & Buckner, 2009). O’Reilly and colleagues used predetermined regions in sensorimotor, motor, visual, auditory, parietal, and prefrontal cerebral cortex to determine functional connectivity to specific cerebellar regions (O'Reilly, Beckman, Tomassini, Ramnani, & Johansen-Berg, 2009). They found parallel circuits in the cerebellum that could be functionally segregated into motor-related and non-motor executive-related areas. Though their results support the notions of parallel cerebro-cerebellar loops, they also find evidence for overlapping sensory and motor input that converge into one cerebellar region, especially at the resolution afforded by functional magnetic resonance imaging (O'Reilly, Beckman, Tomassini, Ramnani, & Johansen-Berg, 2009).

Neuroimaging studies of task-related activation changes have provided another key source of information about the contributions of the cerebellum to cognition. For instance, a seminal study by Petersen and colleagues (Petersen, Fox, Posner, Mintun, & Raichle, 1989) explored changes in blood flow associated with three different language tasks: passive listening or viewing, repeating words, or generating uses for visually or acoustically presented words. As expected from the location of the motor representations in the cerebellum, changes in activity were found for the in the superior anterior cerebellum in the condition that required production. However, a right lateral inferior cerebellar region that was distinct from this region was active in the generate-use versus repetition task comparison (Petersen, Fox, Posner, Mintun, & Raichle, 1989). Since the comparison to the repetition task was meant to account for motor-related aspects of the generate-use task, activation in the lateral cerebellar region was thought to reflect a non-motor contribution of the cerebellum to task performance.

In addition to PET, functional magnetic resonance imaging also emerged as a non-invasive tool to examine system-level changes in brain activity and provided a new means to investigate the role of the cerebellum in cognitive tasks. The cerebellum has now been implicated in a number of cognitive tasks including memory, language, learning, and attention (Fiez, 1996; Thach W. T., 1998; Strick, Dum, & Fiez, 2009). Though many speculations about its role in cognitive tasks have emerged, there is still debate about the exact nature of its function. Given the systematic structure of the cerebellum, many investigators have presumed that it applies the same algorithms across all motor and non-motor tasks. Presumably, distinct subregions within the cerebellum are involved in different tasks through their connections to different cerebral cortical regions.

One obvious contribution of the cerebellum to cognitive processes is its role in articulation. The processes involved in articulatory planning and rehearsal are thought to be a key component of many cognitive tasks (Baddeley, Lewis, & Vallar). Functional neuroimaging studies have consistently shown cerebellar activity in studies investigating overt and covert speech production (Riecker, Kassubek, Groschel, Grodd, & Ackermann, 2006; Soros, Sokoloff, Bose, McIntosh, Graham, & Stuss, 2006; Bohland & Guenther, 2006).

In a series of functional neuroimaging and transmagnetic stimulation studies, Desmond et al. (Desmond, Gabrieli, Wagner, Ginier, & Glover, 1997; Desmond, Chen, & Shieh, 2005) demonstrated the involvement of specific regions of the cerebellum, namely the superior cerebellar hemisphere (VI/Crus I) and the right inferior cerebellar hemisphere (VIIB), in different components of verbal working memory (Desmond, Gabrieli, Wagner, Ginier, & Glover, 1997). They consider these two regions in the cerebellum to work in concert with cerebral cortical areas as part of an articulatory and phonological network involving Broca’s area and inferior parietal cortex, respectively. By comparing intended and actual phonological outputs, the cerebellum is thought to detect discrepancies and update rehearsal commands (Desmond, Gabrieli, Wagner, Ginier, & Glover, 1997). This model supports the idea of distinct cerebellar regions involved in different components of verbal working memory as a function of their connectivity to different cerebral cortical regions that support articulatory or phonological processes.

Within the speech production literature other investigators have studied classic effects attributed to the cerebellum, such as sensorimotor adaptation. Behavioral work demonstrates that sensorimotor adaptation in speech can be attained through a number of manipulations of the speech signal (e.g. delayed auditory feedback, pitch shifts, etc.) that lead to compensation in motor output (Houde & Jordan, 1988; Jones & Munhall, 2000). A neuro-computational model of speech production developed by Guenther et al. (the DIVA model) implements feedforward and feedback mechanisms (Guenther & Ghosh, 2003; Guenther, Ghosh, & Tourville, 2006). The basic premise of this model is based on ideas that are akin to the forward and inverse models proposed in other domains of motor control where motor commands yield sensory predictions. Errors between the expected and sensory are relayed to the premotor cortex to adjust the motor output. The feedforward and feedback systems work in concert to deal with perturbations in speech feedback, which in turn lead to adaptation in speech production. Guenther et al. posit that speech production involves a network of regions with different components of this network involved where premotor cortex, motor cortex, and the cerebellum are involved in feeforward control and the prediction of sensory consequences, whereas superior temporal, inferior parietal cortex, and cerebellum are involved in the feedback control process (Guenther & Ghosh, 2003).

A recent functional neuroimaging study examined the neural basis of these effects using a word production task with either normal auditory feedback or an auditory feedback condition where formants frequencies were shifted in real time (Tourville, Reilly, & Guenther, 2008). They found increases in activity in temporal areas, motor areas, and in the cerebellum. In the DIVA model, errors caused by the unexpected sensory outcome would be represented in the auditory cells in superior temporal cortex, perhaps as increases in activity. The error signals would be used to alter the forward commands provided by motor, premotor, and cerebellar cortex. The results of Tourville et al. were consistent with the predictions of the DIVA model whith increases activity in superior temporal cortex possibly representing the errors and providing the information necessary to correct motor output. However, we note that other models of sensorimotor adaptation presume that the comparison and error between the intended and actual sensory outcome takes place within the cerebellar circuitry, potentially in the inferior olive, and then conveyed by the climbing fibers to the cerebellum.

Although cerebellar damage is commonly associated with disorders of speech production, such as dysarthria, neuropsychological and neuroimaging studies indicate that the cerebellum contributes more broadly to language functions. For example, Fiez et al. (1992) demonstrated in a case study that when damage is limited to the lateral hemispheres, dysarthric symptoms can be absent (Fiez, Petersen, Cheney, & Raichle, 1992). However, upon more detailed investigations of language function, deficits in higher-level language processes were found. A recent study by Stoodley and Schmahmann examined a group of cerebellar patients on phonemic and semantic fluency (Stoodley & Schmahmann, 2009). They also found language deficits that could not be accounted for by mere deficits in language production.

In 2004, a review of 82 functional neuroimaging experiments involving language tasks concluded that higher-level language processes tended to recruit more lateral areas of the right cerebellar hemispheres compared to other language-related tasks that engage articulatory processes (Indefrey & Levelt, 2004). Stoodley et al. conducted a meta-analysis of various processes (motor, sensorimotor, somatosensory, language) and localized linguistic functions to lateral parts of the cerebellar hemisphere in posterior LobeVI and Crus I (Stoodley & Schmahmann, 2009). These findings provide additional evidence against the notion that cerebellar contributions to cognition are limited to articulatory planning.

Despite the advances in identifying and characterizing cerebellar involvement in non-motor processes through anatomical, clinical, and functional neuroimaging studies, a consensus about its functional contributions to cognitive tasks has not been reached. Counfounding factors, that have been attributed to the cerebellum, such as timing, attention, and motor responses, exist in many cognitive tasks and have contributed to the challenges in interpreting its involvement in cognitive tasks. There have also been technical issues that have limited progress. For instance, most studies have not used pulse sequences that afford whole brain coverage and as a result the inferior portion of the cerebellum is often poorly sampled. Moreover, scanner noise that interferes with auditory stimulus presentation, and movement-related artifacts caused by speaking have also hindered advances in studies of language.

1.3 CEREBELLAR INVOLVEMENT IN SPEECH PERCEPTION

Speech perception is an ideal system for investigating cerebellar mechanisms in non-motor processing since it can be influenced by both sensory and cognitive processes without requiring intentional motor output. In addition, various parameters of the speech input can be manipulated independently to address specific questions (e.g. acoustic-spectral, acoustic-temporal, contextual properties, etc.). The number of functional neuroimaging studies of speech perception tasks has markedly risen in the last few years with over 70 published studies generated in a database search since January of 2009. Many different aspects of speech perception have been examined using a wide variety of tasks (e.g., perceptual phonetic categorization, perceptual phonetic discrimination, sentence comprehension, word recognition, etc.).

Though the cerebellum is often excluded from data collection in many speech perception studies, a meta-analysis examining purely auditory tasks identified a region showing the greatest likelihood of activation across studies that extended from Crus I to Lobe VI (Petacchi, Laird, Fox, & Bower, 2005). If closed cerebro-cerebellar loops are recruited for speech perception, knowledge of the cerebral cortical structures involved in normal speech perception could elucidate the role of the cerebellum. In several extensive review articles, Hickok and Poeppel considered findings from neuropsychology and functional neuroimaging to develop and fine-tune a neurally-based framework for auditory speech perception and language comprehension (Hickok & Poeppel, 2007; Hickok & Poeppel, 2004). Their model consists of a dorsal and a ventral stream that contribute to different aspects of speech processing. Both streams diverge from phonological-level (sound-based) representations in the superior temporal sulcus. The dorsal stream provides access to an articulatory network in inferior frontal gyrus and premotor cortex through a sensorimotor interface in temporo-parietal cortex; the ventral stream accesses conceptually-based representations in the anterior temporal lobe through a lexico-semantic interface in middle and inferior temporal gyri.

Anatomical connections between the cerebellum and temporal cortex have not been extensively investigated; however, anatomical studies in non-human primates confirm that the cerebellum receives input from the superior temporal plane and sparse input from the superior temporal sulcus (Schmahmann & Pandya, 1991). Functional connectivity analyses contribute to the existing evidence on the connections between the cerebellum and temporal cortex. Using resting state fMRI, O’Reilly and colleagues show input from auditory cortex and superior temporal cortex to lobes V and VI (O'Reilly, Beckman, Tomassini, Ramnani, & Johansen-Berg, 2009). Using dynamic causal modeling, a method used for determining the direction of the connectivity, Booth et al. found evidence for functional connections between the cerebellum and the lateral temporal cortex that were bidirectional (Booth, Wood, Lu, Houk, & Bitan, 2006).

Speech perception may involve other cerebro-cerebellar loops that do not include the superior temporal cortex, since it is impacted by other sensory modalities and cognitive representations (McGurk & McDonald, 1976; Ganong, 1980). For example, a highly influential psychological model, the motor theory of speech perception presumes that motor representations underlie the basis of speech perception (*for discussion see* Galantucci, Fowler, & Turvey, 2006). A number of studies have shown common networks that are active for speech production and perception. Indeed, a functional neuroimaging study of speech perception even showed activation in ventral premotor cortex during a passive listening task (Wilson, Saygin, Sereno, & Iacoboni, 2004). Though such findings are often used as evidence for the need for motor representation in speech perception, the degree of importance that motor representations have on the actual perception of speech is under considerable debate (*for discussion see* Lotto, Hickok, & Holt, 2009). Changes in activity observed in motor regions need not be essential for speech perception and may instead be a reflection of highly interactive and bidirectional pathways between sensory and motor areas in cortex.

Access to motor representations during speech perception may also be recruited during the acquisition of new phonetic contrasts (Callan, Tajima, Callan, KuboRieko, Masaki, & Akahane-Yamada, 2003; Callan, Jones, Callan, & Akahane-Yamada, 2004; Callan, Callan, Tajima, & Akahane-Yamada, 2006). An fMRI study examining adult language acquisition of a foreign phonetic contrast with perception training (specifically Japanese subjects on the non-native (r/l) contrast) showed greater activity in an entire network of regions involved in speech production (including the cerebellum) after perceptual training (Callan, Tajima, Callan, KuboRieko, Masaki, & Akahane-Yamada, 2003). Furthermore, the changes in activity paralleled subject improvements in performance (Callan, Callan, Tajima, & Akahane-Yamada, 2006; Callan, Jones, Callan, & Akahane-Yamada, 2004). In another study, Callan et al. found similar results for a durational short/long vowel contrast. The results of this study supported the idea that new relationships between articulatory and sensory representations are formed during learning. The authors conclude that the cerebellum may be involved in establishing these new input-output maps through forward and inverse internal models (Callan, Callan, Tajima, & Akahane-Yamada, 2006).

These studies provide a sensorimotor basis for cerebellar involvement in perception. To the extent that interactions between sensory and motor representations may recruit the cerebellum during speech perception adaptation, we take into consideration the location of activation and the relationship between changes in activity and changes in behavior in our regions of interest. Based on a meta-analysis of solely auditory neuroimaging studies, a purely perceptual role for the cerebellum has also been suggested (Petacchi, Laird, Fox, & Bower, 2005). However, information regarding cerebellar involvement in normal speech perception has been underrepresented, and investigations into its specific role in adaptive processes in speech perception have not been conducted. This is surprising given its known contributions to timing and sensorimotor adaptation. We consider the existing evidence on cerebellar involvement in timing mechanisms in speech perception and consider its potential contributions to speech perception adaptation.

1.3.1 Timing in Speech Perception

The cerebellum seems to be recruited for timing mechanisms in both the production and perception of movement (for discussion see Ivry, 1996; Ivry, 2004). The perception of speech critically relies on spectral and temporal properties of the speech signal. Timing information is a key component in distinguishing among several phonetic contrasts (e.g voice onset time). In many languages, there exist linguistic contrasts that differ by only a durational parameter (e.g. vowel length).

Functional neuroimaging studies have investigated the involvement of the cerebellum in discriminating between such durational contrasts. One functional neuroimaging study compared changes in cerebellar activity when participants distinguished between two speech items that differed in either a durational parameter (occlusion duration) or two items that differed in both a durational and a spectral parameter (voice onset time). The findings of Ackermann et al. indicate that a region in the right lateral hemisphere of the cerebellum, Crus I, was more sensitive to the purely durational contrast (Ackermann, Mathiak, Ivry 2004). Similarly, neuropsychological study showed that patients with diffuse cerebellar atrophy had impaired categorical perception for the two items differing in occlusion duration (Ackermann, Graber, et al. 1997), but no difficulties in differentiating items that vary in voice onset time (which may be discriminated using spectral information related to the aspiration) (Ackermann, Mathiak, & Ivry, 2004).

1.3.2 Speech Perception Adaptation

Plasticity in speech perception is evident through our ability to rapidly adapt to speaker variability and noisy environments (Magnuson & Nusbaum, 2007). Many examples of adaptation in speech perception have been formally investigated, including adaptation to foreign-accents and disordered speech, as well as distorted time-compressed and noise-vocoded speech (Altmann & Young, 1993; Fenn, 2003; Maye, Aslin, & Tanenhaus, 2008; Liss, Spitzer, Caviness, & Adler, 2002; Francis, Nusbaum, & Fenn, 2007). However, little is known about the neural substrates that underlie these adaptive processes in speech perception. A few recent functional neuroimaging studies have examined adaptation to time-compressed speech, noise-vocoded speech, and delayed auditory feedback in perception or production tasks (Adank & Devlin, 2009; Toyomura, et al., 2007).

An fMRI study examining adaptation to time-compressed speech showed increases in both the superior temporal gyrus and the inferior frontal gyrus with increasing distortion (Poldrack, et al.). Another study showed decreases in neural activity in the bilateral superior temporal gyri and ventral premotor cortex with adaptation (Adank & Devlin, 2009). No information about cerebellar activation was provided in these studies. It is unclear whether results from the cerebellum were unavailable due to a lack of coverage during the image acquisition or for other reasons. Though inconclusive, the results of these studies suggest that articulatory processes may aid speech perception when the speech signal is degraded. Nevertheless, it is also known that context-dependent information, such as word and sentence identity, can also lead to adaptation in speech perception (Davis, Hervais-Adelman, Taylor, Johnsrude, & McGettigan, 2005; Norris, McQueen, & Cutler, 2003).

Thus, our primary goal is to determine whether the cerebellum is involved in speech perception adaptation and to identify specific regions within cerebellar and cerebral cortex that may be responsible for those adaptive processes. To investigate mechanisms involved in the adaptation *per se,* while avoiding time perception aspects of cerebellar involvement, we employ a speech manipulation that mimics cochlear implants. In this manipulation, the speech distortion affects the spectral but not the temporal parameters of the speech signal. In an attempt to identify the regions that are involved in the adaptation aspect of the task, we will perform correlations between changes in brain activity and changes in behavioral performance. This unique aspect of our study will enable us to determine whether there are non-motor changes in performance that can be attributed to the cerebellum. Finally, functional connectivity analyses will be conducted to give us insight into potential connections between the cerebellum and cerebral cortex involved in normal speech perception and speech perception adaptation.

2.0 A BEHAVIORAL INVESTIGATION OF SPEECH PERCEPTION ADAPTATION TO SPECTRALLY DISTORTED WORDS

In sensorimotor adaptation, the cerebellum may be involved in retuning the correspondence between sensory and motor representations when there is a mismatch between the intended and actual sensory outcome (Wolpert, Miall, & Kawato, 1998). For example, visually-guided reaching tasks, sensorimotor coordination is necessary. If an individual reaches for a target while wearing prism goggles that distort the visual feedback by producing a lateral shift in the visual field, the individual will miss the target. With repeated attempts, the individual corrects their actions to adjust to the distorted visual feedback and eventually reaches the target. The cerebellum plays an essential role in this task; patients with cerebellar damage are unable to adapt to the visual distortion (Pisella, Michel, Grea, Tilikete, Vighetto, & Rossetti, 2004). The role of the cerebellum appears to be even more critical when distortions are introduced gradually (Robertson & Miall, 1999; Ethier, Zee, & Shadmehr, 2008). Greater deficits in a visuomotor adaptation task has been reported when the visual distortion was gradually introduced compared to when there was an abrupt distortion during inactivation of the dentate nucleus in an animal study using monkeys (Robertson & Miall, 1999).

Analogous sensorimotor adaptation to gradually introduced distortions has also been demonstrated in speech production tasks. Studies have examined a range of manipulations including shifts in fundamental frequency (pitch), vowel formant frequency, and timing manipulations of the auditory speech feedback (Jones & Munhall, 2003; Jones & Munhall, 2005; Houde & Jordan, 1988). Participants adjust their motor speech output to compensate for the distorted sensory feedback. In one study, participants produced a completely different vowel sound after they were exposed to gradually introduced vowel formant shifts along the edge of their individual vowel triangle (Houde & Jordan, 1988). A neurally-based computational model of speech production, the DIVA model, accounts for these findings by incorporating feedforward and inverse commands. In their model commands errors derived from sensory predictions are used in a feedback control system to compute the motor commands needed for the the desired sensory output. The cerebellum is a component of both the feedforward control system that maintains the appropriate motor commands and predicts sensory consequences, as well as in the feedback control that corrects the error (Guenther & Ghosh, 2003).

Adaptation also exists in tasks that do not have an intentional motor component. Speech perception, for example, adapts in response to unexpected acoustic speech input. Listeners adapt to highly distorted acoustic speech signals when listening to talkers with strong non-native accents or to patients with dysarthria (Maye, Aslin, & Tanenhaus, 2008; Liss, Spitzer, Caviness, & Adler, 2002). They also adapt to more artificial speech distortions such as synthetic, time-compressed, and noise-vocoded speech (Hervais-Adelman, Davis, Johnsrude, & Carlyon, 2008; Altmann & Young, 1993; Greenspan, Nusbaum, & Pisoni, 1988).

Functional neuroimaging, as discussed in the Introduction chapter, has been used to study the neural basis of speech perception; however, the potential role of the cerebellum to speech perception adaptation has received little attention. As a first step toward developing a neuroimaging study targeting this question, we sought to establish behavioral characteristics of speech perception adaptation. While more naturalistic distortions such as foreign accents have ecological validity, they do not afford tight experimental control over the acoustic-perceptual characteristics of the stimuli. For this reason, much of the past research has used more easily controlled speech distortions such as synthesized speech and time-compressed speech. However, these distortions still have two important limitations. First, they involve manipulations of the temporal characteristics of speech, and thus they might confound together potential contributions of the cerebellum to both timing and adaptive plasticity. Second, it is difficult to incrementally vary the intelligibility of the stimuli using these distortion procedures, and this makes it difficult to implement the types of training paradigms used in sensorimotor and speech production adaptation tasks where distortions can be gradually introduced.

Understanding the adaptation process in speech perception is also beneficial to cochlear-implant patients. Cochlear implants are used for profoundly deaf individuals and create the perception of sound through auditory nerve stimulation in the cochlea. Thus, some of the current literature on speech perception adaptation has been motivated by this field; adaptation to spectrally degraded and shifted speech noise-vocoded speech distortion, which mimic mimic the limited-channel output of chochlear implants in normal-hearing listeners (Shannon, Zeng, Kamath, Wygonski, & Ekelid; Li & Fu, 2006). Speech sounds are separated into a set of band-limited channels through an algorithm that is similar to cochlear implants (Zeng, 2004). The low-frequency temporal dynamics of the original speech are preserved but the spectral resolution within each channel is altered (see Figure 1). Li and Fu used a sophisticated version of this distortion where the frequency channels are also shifted. Intelligibility is greatest for smaller mismatches and declines precipitously as the frequency channels are shifted away from the regions of the spectrum that typify natural speech.

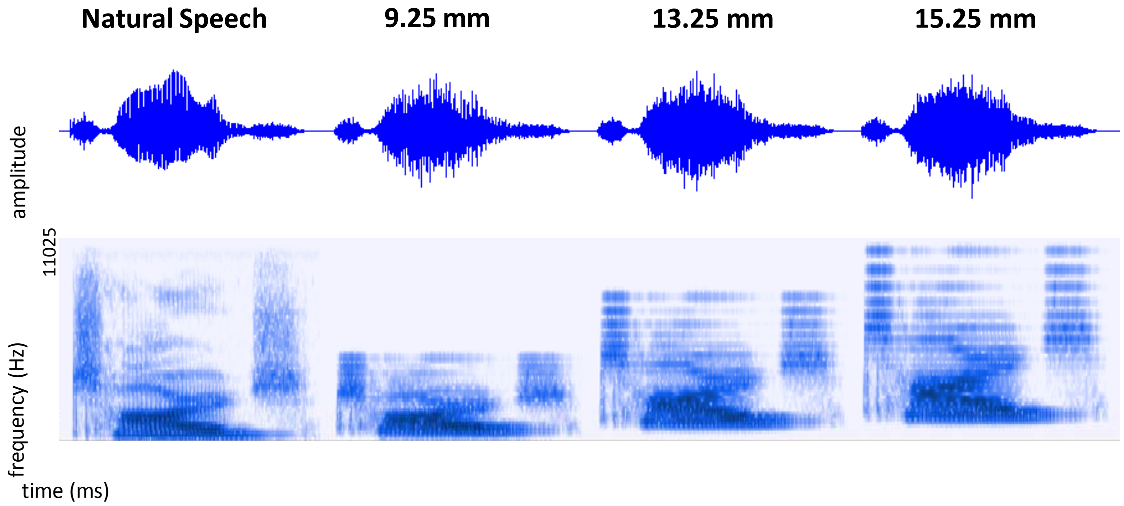


Figure 1. Spectrogram of different distortion levels

The top panel illustrates waveforms (time x amplitude) for a natural utterance “charge” and the same word processed with 9.25, 13.25, and 15.25 mm insertion depths. The temporal envelope of the speech remains quite constant across stimulus distortions. The bottom panel shows the corresponding spectrograms on time by frequency axes, with amplitude shown as dark bands. Note the shifts in the speech spectra. For 15.25 insertion depth, no energy remains below 1000 Hz.

In unpublished work, Holt characterized this relationship by introducing distortions at varying degrees of distortion and examining their effect on intelligibility (Holt, unpublished results). The relationship between the severity of the distortion and intelligibility is demonstrated below (see Figure 2). This procedure enabled us to systematically introduce distortions at varying degrees of intelligibility.

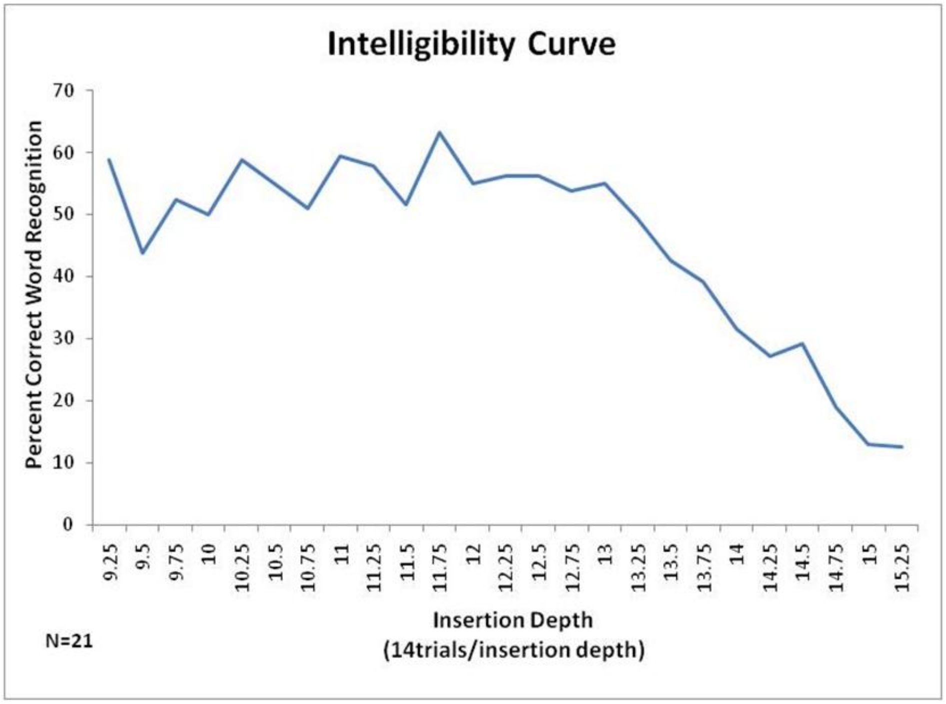


Figure 2. Intelligibility curve for spectral distortion

Percent word correct is plotted as a function of distortion level. Twenty-one undergraduate-aged participants from the greater Pittsburgh area were randomly presented with stimuli at different distortion levels. Each distortion was presented 14 times for a total of 350 trials.

Fu et al. examined the effects of distortion severity of noise-vocoded speech on adaptation by using two different distortion severities (Li & Fu, 2006). They found that improvements in performance depended on the severity of the distortion. Though the literature on adaptation in speech perception is growing, many studies employ similar training conditions to achieve learning; These paradigms typically include giving explicit feedback during training to provide the information necessary to drive learning. For example, significant improvements in performance are often observed after a training condition that pairs orthographic feedback (written word) with the auditory stimulus (Fenn, 2003). However, it is known that adaptation of speech sounds can also be achieved when no explicit feedback is provided about the identity of the sound. For example, ambiguous sounds presented within the context of a word, can tune the categorization of that sound (Norris, McQueen, & Cutler, 2003).

Davis et al. demonstrated that contextual information can also affect adaptation for noise-vocoded speech. In one study they showed that performance was improved when a distorted word was presented within the context of a sentence (Davis, Hervais-Adelman, Taylor, Johnsrude, & McGettigan, 2005). When participants were trained on nonword sentences, no improvements in performance were observed. These results suggest that higher level contextual information provided by the sentence can be used to drive learning for noise-vocoded speech distortions.

An important aspect of cerebellar mechanisms in sensorimotor adaptation is that the error signals guiding the adaptation can be generated internally through a comparison between a predicted sensory consequence and the actual sensory outcome, without the need for explicit feedback. This view is supported by studies demonstrating that the cerebellum is also involved in adaptation when the distortion is imperceptible and gradually introduced (Yamamoto, Hoffman, & Strick, 2006). Therefore, this manipulation allows us to gradually increase the distorted sensory feedback during training and allow word recognition, which can be used to drive learning.

In speech perception, higher level contextual information (such as the identity of the word or the meaning of a sentence) can guide the selection of prelexical phonological representations and be compared to the distorted speech input. Errors between prelexical phonological and sensory information can be generated and used to guide the adaptation. Hence, the ability to access lexico-semantic or conceptual representations could impact the degree to which adaptation occurs when explicit feedback is absent. Therefore, in the first experiment we examine the role of distortion severity and explicit feedback on speech perception adaptation. Specifically, we compare adaptation gains for conditions which differ in whether or not they provide explicit orthographic feedback or sequentially increment the distortion levels without explicit feedback. In Experiment 2, we use an intermediate distortion level to test the hypothesis that increasing the intelligibility of the distortion level allows lexically-driven adaptation to occur even in the absence of explicit feedback.

2.1 EXPERIMENT 1

In this experiment, we explore the importance of distortion severity and orthorgraphic feedback on adaptation to spectrally distorted speech. In the sensorimotor adaptation, cerebellar mechanisms are thought to be engaged when learning results from an internally-generated error signal caused by a mismatch between motor output and sensory input. The role of the cerebellum is believed to be even more important when the distortion is gradually introduced. Thus, we examined whether the severity of the distortion affects adaptation in speech perception by comparing training with a severe distortion versus sequentially presented, incremental increases in the distortion during training. A majority of the studies investigating adaptation in speech perception use some form of explicit feedback. However, cerebellar mechanisms have been associated with more implicit learning processes since improvements can occur when the error signals are likely to be imperceptible. Thus, in a between-subject design, we compared improvements in word recognition after with and without orthographic feedback.

2.1.1 Methods

2.1.1.1 Participants

Ninety-two participants were paid eight dollars for their participation in this study and gave informed consent prior to participation in the study. They were required to be normal hearing and native listeners of English. A total of eleven participants were eliminated from the analysis because they failed to follow instructions or meet the prerequisite exclusion criteria. A total of four participants were excluded from the analysis due to technical errors. The data from the remaining seventy-seven participants were included in the analysis. The participants were distributed as follows severe distortion with feedback, N= 21; severe distortion with no feedback, N=17; incremental distortion with feedback, N=18; and incremental distortions without feedback, N=21.

2.1.1.2 Stimuli

The words used in all of the experiments described below were drawn from Egan (1948; lists 1-9). Experiments 1-3 used lists 1-9. Each list was composed of 50 phonetically-balanced English monosyllabic words. Stimuli that were pronouns or plurals were removed from the lists. A female speaker (LLH) uttered each word into an Electrovoice RE 20 microphone connected to a digital Marantz PMC670 recorder,with 16-bit resolution at a sampling rate of 22050 Hz. Stimuli were saved as individual *.wav files, and equated across all lists for RMS (root mean square) amplitude. We used Tiger Speech to introduce spectral shifts in the stimuli (http://www.tigerspeech.com/tst_tigercis.html) based on the methods of Fu & Galvin (2003) and Li and Fu (2006) for a 20-Channel sine-wave-vocoded speech Analysis Filter Design (Analysis_NU24_20.txt), 200-7000Hz, 24 dB/oct Carrier Filter Design (), 200-7000 Hz, 24 dB/oct).

Each speech token was band-pass filtered into 20 frequency bands using eigth-order Butterworth filters. The frequency characteristics of the bands were based on the Nucleus-24M cochlea implant (116-7871 Hz). Each band was half-wave rectified to extract the temporal envelope and low-pass filtered at 160 Hz. This envelope served to modulate a carrier band. The following equation was used to calculate the frequency range of the carrier band: $p(i) = P0 + 0.75 * i$, $i=0, 1, 2, 3., 20$. The carrier band frequencies assumed 35-mm cochlea (Greenwood, 1990) and were calculated using the following equation: $f(i)=165.4 * (10P(i) * 0.06 - 0.88)$. The two equations were combined to determine the corner frequencies of carrier bands for a given insertion depth. Cross-over attenuation between adjacent bands was -3 dB. The carrier bands were summed to create the spectrally-shifted speech tokens.

For the purposes of the present experiment, P0 varied between 9.25 and 15.25 mm from the apex of the cochlea, simulating a range of shallow to deep insertion depths and shifting speech spectra incrementally upward on the frequency dimension. Carrier bands incrementing in steps of 0.25 mm insertion depth were calculated between these endpoint values. For example, at an insertion depth of 9.25 mm speech was shifted upward in frequency such that there was no spectral energy below 448 Hz. At the most severe distortion, 15.25 mm, there was no spectral energy below 1214 Hz. These signal distortions create a complex mapping challenge for word recognition since a great deal of information in the speech signal is carried below 2000 Hz.

2.1.1.3 Procedure

While seated in sound-attenuated booths, participants heard the distorted speech stimuli through Beyer DT-150 headphones at approximately 70 dB. Participants performed a free word recognition task by typing their response on the computer keyboard. All participants performed a 50 trial pretest and a 50 trial posttest with high distortion stimuli (15.25 insertion depth) that were randomly selected from the total set of words. Each trial encountered during pretest, training, and posttest was a new stimulus, which eliminated the possibility for participants to learn the mapping between a specific item and the distorted speech input.

Training Conditions

Severe Distortion/Feedback Training. Participants were presented with 250 trials of distorted words that were randomly selected from the total set of words. On each trial, they were presented with a word through headphones that was distorted at a 15.25 insertion depth. Following their response, they were presented with the printed form of the word, which appeared in the middle of the computer screen, paired with a repeat of the same auditory speech stimulus. The feedback was based on previous studies (Fenn, Nusbaum, Margoliash, 2003)**.**

Incremental Distortions/Feedback Training Condition. Participants were presented with 250 trials of distorted words that were randomly selected from the total set of words. After each 10 trials, the difficulty of the distortion was increased by an insertion depth of .25 beginning at an insertion depth of 9.25 and ending with an insertion depth of 15.25. Following their response, they were presented with the printed form of the word paired with a repeat of the auditory stimulus. The feedback was based on previous studies (Fenn, Nusbaum, Margoliash, 2003)**.**

Incremental Distortion/No Feedback Training. Participants were presented with 250 trials of distorted words that were randomly selected from the total set of words. After each 10 trials, the difficulty of the distortion was increased by an insertion depth of .25 beginning at an insertion depth of 9.25 and ending with an insertion depth of 15.25. Participants heard only 10 trials at the 15.25 insertion depth during training. No feedback was provided for this condition.

Severe Distortion/ No Feedback Training. Participants were presented with 250 trials of distorted words that were randomly selected from the total set of words. On each trial, they were presented with a word through headphones that was distorted at a 15.25 insertion depth. No feedback was provided for this condition.

2.1.2 Results

In our first analysis, we examined the effect of Feedback in each distortion condition (see Table 1). In the severe distortion condition, an independent T-test with effect (Posttest-Pretest) as the dependent measure revealed a significant difference between the Feedback and No Feedback condition T(36) = 2.8, p = .008. In the incremental distortion condition, no significant effect of feedback was found. Whereas explicit feedback was necessary for learning to occur with training at the most severe levels of acoustic distortion, incremental introduction of the distortion led to learning even without explicit feedback.

Table 1. Experiment 1. Word recognition accuracy

Percent accuracy on word recognition in the pretest and posttest conditions; standard error of the mean values are reported in parentheses

Training Condition		Pretest	Posttest	(Posttest – Pretest)
Severe Distortion	Feedback	7.4 (1.7)	23.8 (3.1)	16.4 (2.3)
	No Feedback	5.4 (1.2)	12.8 (2.2)	7.4 (1.0)
Incremental Distortions	Feedback	6.4 (1.5)	21.2 (2.7)	14.8 (2.3)
	No Feedback	6.8 (1.0)	18.9 (2.9)	12.2 (2.6)

To compare the effects of Feedback on adaptation directly with the effects of gradually introduced distortions on adaptation, we contrasted the Feedback condition at the severe distortion level to the No Feedback condition with the incremental distortions. Again, no significant effects were found. Word recognition for severely distorted speech improved just as much among the group of participants exposed to incremental acoustic distortions without feedback during training, as among participants explicitly trained with the severely distorted speech with explicit feedback during training.

If improvements in performance on the distorted speech are lexically-mediated, then we should see some evidence for correlations between performance during training and improvements in performance from pretest to posttest. Thus, we used a Pearson correlation to determine whether performance during training correlated with lexico-semantic improvements in performance. Indeed, all training conditions showed a significant correlation with the difference between performance on the posttest condition and the pretest condition, R-value > 0.5, $p < 0.05$.

2.1.3 Discussion

We found that the degree of perceptual learning elicited by feedback-driven training is equivalent to a training paradigm with no explicit feedback, but for which the severity of the acoustic distortion is incremented gradually. The fact that each stimulus encountered by the participants was a novel word suggests that the adaptation is occurring between the distorted acoustic input and the prelexical phonological representations. Thus, access to lexico-semantic information with less severe distortions must be providing the necessary information to successfully drive adaptation at the prelexical phonological level. This is consistent with theoretical models of speech perception that have incorporated such processes, whereby lexico-semantic information interacts with prelexical phonological representations to influence learning. In contrast, exposure to severe acoustic distortions without external feedback does not result in comparable degrees of adaptation, which suggests that lexico-semantic information is not available to influence learning. If access to lexico-semantic information is sufficient to drive learning, perhaps training on a more intermediate distortion, where lexico-semantic information can be accessed some of the time, would also result in significant improvements in performance even when no explicit feedback is provided during training. Our next experiment is aimed at addressing this question.

2.2 EXPERIMENT 2

Our motivation for Experiment 2 is to develop a paradigm that is suitable for a functional magnetic resonance imaging experiment. In Experiment 1, we found equivalent performance improvements to distorted speech with feedback training and in a training condition without feedback but where the acoustic distortion was incremented gradually. However, in a functional neuroimaging design, incrementally increasing the distortion would confound together improvements in performance and changes in the acoustic signal during training; hence maintaining the same distortion level throughout training would be more desirable. In addition, if we achieve a slower rate of learning with the intermediate distortion, this would be beneficial for examining changes in brain activity during training. Therefore, we examined an intermediate distortion level where access to lexico-semantic representations is achieved 40-50% of the time with no feedback to see whether we still get significant pretest to posttest learning.

2.2.1 Methods

2.2.1.1 Participants

Eight, undergraduate-aged volunteers were paid eight dollars for their participation in this study. All participants gave informed consent, according to procedures established by the Institutional Review Board at Carnegie Mellon University.

2.2.1.2 Stimuli

The stimuli (13.25 and 15.25 insertion depths) were drawn from the same list of words as experiment in Experiment 1.

2.2.1.3 Procedure

While seated in sound-attenuated booths, participants heard stimuli through Beyer DT-150 headphones at approximately 70 dB. Participants performed a free word recognition task by typing their response on a keyboard. Participants were presented with normal and distorted words. There were five different condition blocks. The first block of 30 trials was a control condition with Normal words that were not distorted. The second block was of 30 trials made up the pretest consisting of words, all at the high distortion level (15.25 insertion depth). The third block consisted of 60 trials at the intermediate distortion level (13.25 insertion depth) and made up the training blocks. The fourth block of 60 trials was a posttest (15.25 insertion depth) and made up the posttest. And the fifth and final block of 60 trials was a second posttest (13.25 insertion depth) and was originally aimed to determine retention of the adaptation after exposure to a different level distortion.

2.2.2 Results

Mean performance on normal words was 92 (+/-) 2 percent accuracy on whole word recognition. A significant improvement in word recognition performance from pretest to posttest for the severe distortion condition T(7) = 2.8, p=0.03. Note that the posttest condition consisted of 60 trials whereas the pretest condition had 30 trials. Equivalent results were obtained when the analysis was restricted to the first 30 trials of the posttest condition. Participants improved in their performance during training, with an average of 37% (+/-) 6% in the first 15 trials of training, to 55% (+/-) 4% in the last 15 trials of training, a significant difference of $p < .05$.

Table 2. Experiment 2. Word recognition accuracy

Percent accuracy on word recognition in the pretest and posttest conditions; standard error of the mean values are reported in parentheses

Training Condition	Pretest	Posttest	(Posttest – Pretest)
Intermediate Distortion (13.25)	6.3 (3)	15.8 (3)	9.5 (3)

2.2.3 Discussion

We have demonstrated performance improvements in a free word recognition task after training on an intermediate speech distortion without explicit feedback. Figure 2 shows a recognition rate of about 40-50% for the 13.25 insertion depth compared to the 10% observed for the 15.25 insertion depth. In the absence of explicit feedback, adaptation must rely on internally-generated feedback. Accurate lexico-semantic access may guide the adaptation process by activating the correct prelexical phonological representations, which can then be used to determine the error between the actual and desired sensory input. Thus, the ability to access lexico-semantic representations could affect the ability to generate the internal error signals between prelexical phonological and lexico-semantic representations. A 40% compared to a 10% success rate in lexico-semantic access can lead to significant improvements in the ability to drive the adaptation.

2.3 CONCLUSIONS

During a free word recognition task of spectrally distorted speech, we examined the effects of distortion severity and feedback on adaptation. We obtained comparable improvements in behavioral performance for conditions when explicit feedback was provided during training compared to when the distortion during training was gradually incremented. However, training on the severe distortion without explicit feedback did not produce similar gains in performance. Figure 2 demonstrates the relationship between distortion levels and intelligibility. When the distortion level is at the 15.25 insertion depth, word recognition only occurs 10% of the time. Training with this distortion level leads to poor lexico-semantic access and negatively impacts adaptation. Importantly, improvements in performance generalize to new stimuli that are not encountered during training, which suggests that the adaptation is not of the lexico-semantic representations, but that it is occurring at a prelexical phonological level of representation.

In a separate experiment we reduced the distortion level during training, without explicit feedback, to 13.25 insertion depth (which according to Figure 2) has a word recognition rate of 40-50%. This condition also showed significant gains in adaptation. Taken together, these findings support theoretically-based ideas in speech perception that emphasize lexically-mediated adaptation of prelexical phonological representations (McClelland, Mirman, & Holt, 2006; McQueen, Norris, & Cutler, 2006).

One question that is unresolved is whether speech perception adaptation is accomplished by mechanisms similar to those demonstrated for sensorimotor adaptation and, in particular, whether the cerebellum plays a role. In the sensorimotor domain, sensory consequences can be predicted from motor commands. When the intended and actual sensory outcome are different, error signals can be generated. The resulting errors can be used to generate new motor commands (Ito, 2008). Similarly, in speech perception adaptation, interactions between lexico-semantic and prelexical phonological representations may help in generating internal error signals when the distorted speech input does not map onto prelexical phonological representations. In this instance, error signals could be used to adjust the perception of the mapping between the distorted speech input and prelexical phonological representations. To gain some insight into this question, we examine the neural basis of these effects in Chapter 3.

Acknowledgements

I would like to thank Christi Gomez, Anthony Kelly, Rob Schwartz, and Marge Gibson for their help in data collection and in checking for spelling errors in participant responses.

3.0 INVESTIGATING CEREBRAL AND CEREBELLAR INVOLVEMENT IN SPEECH PERCEPTION ADAPTATION USING FMRI

There has been considerable evidence suggesting cerebellar involvement in non-motor functions, including speech perception, emotion, cognition, and attention (Daum & Ackermann, 1995; Thach W. , 2007; Strick, Dum, & Fiez, 2009). Given the uniform microstructure of cerebellar cellular organization, its function is thought to be similar across motor and non-motor tasks. One role that has been attributed to the cerebellum in sensorimotor adaptation is retuning relationships between sensory and motor representations based on prediction errors (Schultz & Dickinson, 2000; Wolpert, Miall, & Kawato, 1998; Doya, 2000). Computational models use forward models to predict the sensory consequence from a motor command. When there is a discrepancy between the predicted and actual sensory outcome, an inverse model computes the necessary motor command to achieve the desired outcome (Wolpert, Miall, & Kawato, 1998).

Though cerebellar involvement in non-motor processes has become more widely accepted, its involvement in the adaptation of non-motor functions is still unclear. Recently, however, the idea of internal models has been extended to non-motor tasks (Ito, 2008). Ito suggests that that the role of the cerebellum during cognitive functions may be similar to its role in sensorimotor adaptation, with forward and inverse models working together to solve problems. Ito proposes that mental representation for cognitive tasks can be tuned through two mechanisms: 1) error signals that arise from comparisons between the outcome of a forward model and the actual outcome (predicted vs. actual sensory outcome comparison), and 2) feedback-error signals (actual vs. desired outcome comparison) derived from an inverse model that does not successfully achieve the desired outcome (Ito, 2008).

Speech perception provides an ideal system for investigating cerebellar involvement in non-motor adaptive processes because it does not require an intentional motor component. During speech perception, listeners must dynamically adapt their perception to accommodate for the irregularities in the speech signal, adjust to idiosyncrasies of different speakers, and become accustomed to different dialects and foreign accents. Adaptation in speech perception has been demonstrated for natural manipulations such as foreign accents, gender, and speech disorders, as well as more synthetic distortions of the speech signal that can be more carefully manipulated to address specific questions. A number of behavioral studies have successfully demonstrated adaptation to synthetic speech, time-compressed speech, and noise-vocoded speech distortions after training with explicit feedback (Altmann & Young, 1993; Davis, Hervais-Adelman, Taylor, Johnsrude, & McGettigan, 2005; Greenspan, Nusbaum, & Pisoni, 1988).

In a series of experiments in Chapter 2, we showed that perceptual adaptation to spectrally distorted words was achieved without feedback by gradually introducing the severity of the speech distortion or by using an intermediate distortion level. Importantly, adaptation generalized to new word stimuli that were not used during training. Based on these and other findings, we concluded that adaptation was driven by lexico-semantic information (word knowledge). When a lexico-semantic word is chosen, prelexical phonological representations are activated. Analogous to errors arising from sensory and motor comparisons in the sensorimotor literature, errors in speech perception could potentially arise from comparisons made between the activated prelexical phonological representation and the distorted acoustic input; error information could be used to retune the mapping between acoustic and prelexical phonological representations. Thus, adaptation can result even when incorrect words that share prelexical phonological components are chosen.

Despite a rich behavioral literature exploring adaptation to different speech distortions, little is known about the neural mechanisms that underlie the adaptive processes in speech perception. Since the cerebellum is involved in sensorimotor adaptation and has been recently implicated in performing similar adaptive functions in non-motor tasks, we were especially interested in whether we could find evidence for its involvement in speech perception adaptation. The functionally distinct and closed cerebro-cerebellar loops found to underlie cerebellar interactions with cortexsuggest that knowledge of the cerebral cortical structures involved in normal speech perception can be very informative in forming predictions about the involvement of the cerebellum and cerebral cortical structures in speech perception adaptation.

A neurally-based model of language comprehension put forth by Hickok and Poeppel provides a framework for considering contributions of different cerebral cortical structures to speech perception (Hickok & Poeppel, 2007). According to this model, a phonological network is accessed through representations in the posterior superior temporal sulcus and diverges into two processing streams. In the dorsal stream, it interacts with a temporoparietal cortical region in the superior temporal plane (area Spt) that provides access to sensorimotor motor representations. These sensorimotor representations, in turn interface with an articulatory network in the inferior frontal gyrus and the anterior insula. In this framework, the phonological network in the superior temporal sulcus also interacts with a lexico-semantic interface in middle and inferior temporal cortex in a ventral stream of information processing. Thus the representations engaged during speech perception adaptation may recruit specific cerebral cortical structures based on the type of information that is involved in prediction error and the type of representations that are altered.

Our behavioral results from Chapter 2 suggest that adaptation to spectrally distorted speech may involve a remapping of acoustic input onto the prelexical phonological representations. This could occur directly or indirectly through intermediate representations which then select the appropriate prelexical phonological information. Therefore, cerebral cortical regions that have been implicated in prelexical phonological processing, such as superior temporal cortex, may show changes in neural responses before and after training that parallel improvements in partial word accuracy. Regions involved in lexico-semantic access, such as inferior or middle temporal cortex, may be more likely to show a pattern during training that predicts behavioral improvements in performance.

Recent functional neuroimaging studies have examined neural responses to distorted, time-compressed speech. One study showed increases in BOLD responses to time-compressed speech stimuli in the superior temporal gyrus and the inferior frontal gyrus (Poldrack, et al., 2008). In another study that examined time-compressed speech, initial presentation of the speech distortion also led to increases in the superior temporal gyrus as well as in the ventral premotor gyrus. Adaptation to the speech distortion subsequently decreased activity in these regions (Adank & Devlin, 2009). The authors proposed that adaptation in the perception of time-compressed speech likely results from the interactions between motor and sensory representations, which can aid speech perception when speech input is distorted.

To our knowledge, studies examining speech perception adaptation have yet to explore cerebellar involvement in these tasks. This is surprising given the well-established involvement of the cerebellum in sensorimotor adaptation, as well as its known contributions to specific aspects of speech perception including phonetic discrimination (durational contrasts), phonetic learning, and passive listening (Ackermann, Graber, Hertrich, & Daum, 1997; Mathiak K. , Hertrich, Grodd, & Ackermann, 2002; Callan, Tajima, Callan, KuboRieko, Masaki, & Akahane-Yamada, 2003). Furthermore, theoretical models of speech perception emphasize the role of lexico-semantic representations in speech perception and speech perception adaptation (McClelland, Mirman, & Holt, 2006; Norris, McQueen, & Cutler, 2003). Thus our goal for this study was to begin to close the gap between our knowledge of behavioral effects in speech perception adaptation and the neural systems that underlie these processes by using a functional magnetic resonance imaging paradigm.

3.1 METHODS

3.1.1 Participants

Twenty-three healthy volunteers, all right-handed, participated in this study; five participants were not included in the analysis due to excessive head motion, one was eliminated due to equipment malfunction, and two were removed due to incidental neurological findings in the cerebellum. The remaining participants were used in the group analysis (6 women, 9 men; mean age 23.3 +/- 0.8). The subjects provided informed consent prior to their participation according to the University of Pittsburgh Institutional Review Board and were paid $60 upon completion of the study. Upon careful examination of individual results, we found compromised results (extremely low mean signal intensity) in our cerebellar region of interest in one individual. Thus, our subsequent data analysis including graphs and correlations for all regions did not include the results from this participant.

3.1.2 Materials

Stimulus presentation was controlled using “E-prime” Software (Schneider 2002). The stimuli were taken from the experiments in Chapter 2. In the Pretest and Posttest conditions, we used a 15.25 insertion depth. In the training condition, we used words with a 13.25 insertion depth. This procedure described below is a slow event-related design based on Experiment 2 in Chapter 2.

3.1.3 Experimental Procedure

The participants were presented six 11-minute runs consisting of 30 trials. On each trial, the participant was presented with an acoustic stimulus through headphones and asked to guess the identity of the word. There were two different response conditions: written response and no response. On eighty percent of the trials, participants were asked to write their response down on a note card and turn the card, whereas on the other twenty percent of the trials, participants were asked to think of the identity of the word and remain still. The response condition was pseudo-randomly selected with the constraint that half of each response type occurred during the first 15 trials and the other half during the second 15 trials. The six runs comprised different speech conditions. The first run was always normal speech, the second run a pretest (stimuli were of high distortion level words, insertion depth 15.25), the 3rd and 4th runs made up the training condition (stimuli were of intermediate distortion level, insertion depth 13.25), and the 5th run was the posttest (stimuli were of high distortion level words, insertion depth 15.25). The 6th run was made up of additional posttest trials, but was not included in the analyses due to the fact that the majority of the participants showed movement beyond 5 mm in any given direction during this final run.

At the beginning of each trial, a fixation cross was presented on the screen during stimulus presentation. After the first 2 seconds, the participant was prompted with a cue that indicated whether or not they should write down a response. A question mark was presented to indicate a written response, and an “X” was presented to indicate no response. The duration of the response period in both cases was 8 seconds and was followed by a resting period that lasted 12 seconds. The fixation period was marked with a red fixation cross. During the response period, participants wrote their responses on a set of note cards with a pen. Each response was written on a new note card. Participants were told to turn the note card immediately after they wrote down a response or immediately upon seeing the red fixation cross, whichever came first. For each run, a new set of cards was placed in the participants’ hands.

3.1.4 Data Acquisition

Subjects were scanned using a 3.0 Tesla Siemens Allegra, Scanner. Structural images were collected using a T2-weighted pulse sequence in 38 contiguous oblique slices (3.2 mm x 3.2 mm x 3.125 mm) parallel to the AC-PC line. The AC-PC slice was selected for each individual to allow for maximum coverage of the cerebellum while ensuring coverage of temporal and parietal cortex. Thirty-eight functional slices were collected in the same location as the structural slices using a one-shot echo-planar imaging (EPI) pulse sequence [epmax64] (TR = 2 sec, TE = 25, FOV = 200 mm, flip angle = 70 °). Sagittal high-resolution T1-weighted MP- images (1 mm x 1mm x 1mm) were also collected at the beginning of each scan session.

3.1.5 Data Analysis

3.1.5.1 Behavioral Analysis

Each response written response was retyped into a spreadsheet. Each response was then phonetically coded using International Phonetic Alphabet (IPA) conventions by a phonetically-trained linguist. The coded responses were then entered into a custom-designed program that computed the phoneme accuracy of each response. In this software algorithm, the first phoneme in the response was labeled correct if it also was found in the first position of the stimulus. If the first phoneme was incorrect, it was compared to the second position and so on until a match was found. Once it was matched, the second phoneme in the response was compared to the next position in the stimulus until a match was found. If no match was found for the first phoneme, it was labeled incorrect, and the same procedure was applied to the ensuing phonemes. The total number of phonemes correct was multiplied by the ratio of the stimulus/response number of phonemes or the response/stimulus number of phonemes, according to length. The numerator was always the shorter, and the denominator the longer of the two to determine partial word accuracy scores. A two-tailed T-test was performed to determine whether there were significant differences in behavioral performance between the pretest condition and the posttest condition.

3.1.5.2 Imaging Analysis

The data were analyzed using Neuroimaging Software Package (NIS 3.6), developed at the University of Pittsburgh and Princeton University. Automated Image Registration (AIR 3.08) was used to reconstruct and correct for subject motion (Woods, Cherry, & Mazziotta, 1992). Images with movement beyond 4 mm or 4 degrees in any direction were excluded from the analysis. The images were then detrended to adjust for scanner drift within runs. The skulls were then removed from each structural image and the remaining brain tissue was co-registered to a common reference brain which that was chosen from among the subjects in the dataset (Woods, Mazziotta, & Cherry, 1993). Functional images were transformed into the same space by scaling the images to a global mean and smoothed using a three-dimensional Gaussian filter (8mm FWHM). The reference brain was then transformed to Talairach space using AFNI (Talairach & Tournoux, 1988).

Group Analysis.

Pretest vs. Posttest contrast. A voxel-wise repeated measures ANOVA was conducted on the fMRI data using subjects as a random factor. In a two-way ANOVA, condition (Pretest, Posttest) and scan (each 2 second TR, 11 TRs per trial) were used as within-subject factors. The entire 22 seconds of each trial was used for analysis. The first trial of each run was removed from the analysis due to the fact that subjects often reported not being ready for the first stimulus presentation.

Individual Analysis.

To identify loci of cerebellar activity within individual subjects, voxel-wise ANOVAs were performed on the detrended images using trial as a random factor and condition (Pretest, Posttest) and scan time (each 2 second TR, 11 TRs per trial) as within trial factors. Each subject's MPRAGE, T-1 weighted high resolution structural brain was aligned to the functional image by using the "align.py" AFNI script The activation cluster maps were then visually displayed onto the high resolution structural images using a threshold p-value < 0.05 (N=12), or p <0.01 (N=2).

Timecourses of activation. The activation clusters identified through the interaction between condition (pretest, posttest) and time were used to define regions of interest (ROIs) for more in-depth analyses. Timecourse information for each ROI was extracted and a percent signal change for each timepoint in each ROI was calculated using the first and last TRs as the baseline.

Correlations.

Using SPSS, we performed several Pearson correlations on the timecourse data extracted from our ROIs. One analysis was aimed at assessing the relationship between changes in brain activity during training and improvements in behavioral performance. For this measure, for each ROI, the average percent signal change across all TRs during Training was correlated with improvements in performance. Performance improvements were measured as the difference between partial word accuracy during Posttest versus the partial word accuracy during Pretest. A second measure was used to assess whether any differences in brain activity between Pretest and Posttest reflected the changes in behavioral performance from pretest to posttest. For this measure (The average % signal change across all TR's during the Posttest – The average % signal change across all TR's during the Pretest) was correlated with improvements in performance measured as (Partial word accuracy during Posttest – Partial word accuracy during Pretest).

3.2 RESULTS

3.2.1 Behavioral Results

During the normal speech condition, participants identified 76 (+/-) 3 % of the words correctly. Since the words presented in this condition were clearly-uttered instances of relatively high-frequency English words, scanner noise appears to have affected participant word recognition. The same influence seems to have persisted for blocks of trials with distorted speech. Therefore, we used a finer-grained analysis of behavior and measured partial word accuracy scores. A measure of partial word accuracy should reflect changes in the accuracy of the prelexical phonological representations; presumably, with sufficient training accuracy improvements at the sublexical level would lead to improvements at the whole word level as well. Significant differences in behavioral performance were found in a comparison between the pretest and posttest condition, T(15) = 3.6, p = 0.003 (see Table 3).

Table 3. Partial word accuracy

Mean percent partial word accuracy for each condition. The standard errors of the mean are reported in parentheses.

	Normal	Pretest 15.25	Training 13.25	Posttest 15.25
Partial word accuracy	91 (1)	15 (1)	29 (2)	21 (1)

Significant differences in behavioral performance were found in a comparison between the pretest and posttest condition, T(15) = 3.6, p = 0.003.

3.2.2 fMRI Results

We first examined the contrast between our Pretest and Posttest condition. Previous studies in the sensorimotor domain have highlighted the need to account for differences in sensory, motor, and cognitive demands of an adaptation task in order to isolate the regions associated with the adaptation (Clower, Hoffman, Votaw, Faber, Woods, & Alexander, 1996). Therefore, this was the best reflection of brain regions involved in the adaptation process. This is especially important in our study given the extreme nature of the acoustic distortion.

3.2.2.1 Pretest vs. Posttest Contrast

A voxel-wise ANOVA using subject as a random factor and condition (Pretest, Posttest) and scan time (2 second-TR, 11 TRs per trial) as within subject factors was conducted. For the condition by time interaction, the resulting activation clusters that exceed a significance threshold of $p < 0.001$ and a contiguity threshold of 5 voxels are listed in Table 4 and shown in Figure 3.

Table 4. (Pretest, Posttest) x scan time (T1-T11)

Regions displaying a condition x time interaction (p<0.001)
The reported location of the regions were identified using Talairach atlas **(Talairach & Tournoux, 1988)**. *The location of the cerebellar regions were identified using the Schmahmann Atlas coordinates **(Schmahmann, et al., 1999)**.

	Region of activation	BA	Size (#voxels)	Peak Talairach Coordinates (x, y, z)
Motor and Premotor Cortex	Left Postcentral gyrus		93	-42, -36, 56
	Right Postcentral gyrus		7	38, -25, 54
	Left Postcentral gyrus		5	-39, -34, 49
	Right Precentral gyrus	6	5	42, 0, 34
	Left Precentral gyrus	6	8	-47, -5, 32
Frontal Cortex	Left superior frontal gyrus	8/9	25	-5, 33, 45
	Left middle frontal Gyrus	8	9	-32, 41, 45
	Right superior frontal gyrus	8/9	6	19, 54, 37
	Left medial Frontal Gyrus	9/10	12	-6, 56, 14
	Right Middle Frontal Gyrus	10/46	19	42, 48, 16
	Right middle frontal gyrus	9	23	37, 22, 30
Inferior Frontal and Insular Cortex	Left inferior frontal gyrus	44/45	7	-50, 12, 11
	Left insula	13	5	-29, 19, 7
	Right insula	13	18	35, 17, 6
Temporal Cortex	Right superior temporal sulcus	41/22	38	44, -38, 6
	Left posterior inferior/middle temporal gyrus	20/37	6	-64, -45, -8
Occipital Cortex	Right middle occipital gyrus	19	6	48, -77, 27
Cingulate	Left cingulate	31	12	-21, -28, 35
Cerebellar Cortex*	LobeVI		22	-21, -54, -15
	Lobe V/VI		13	23, -35, -25
	Left lobe VI/Crus I		12	-40, -42, -28
	Right Crus I		5	34, -46, -38

3.2.2.1 Correlation Analysis

A subset of these regions identified through the PreTest vs. Posttest contrast showed significant correlations with our behavioral measures or with the BOLD signal during training, namely the right superior temporal sulcus and right Crus I. We examined %BOLD signal change during training as well as the %BOLD signal change **difference** between the Pretest and the Posttest. We found a significant correlation between %BOLD signal change in right Crus I during training and the improvements in behavioral performance measured as the difference in partial word accuracy from Pretest to Posttest. We also found a significant correlation between the **difference** in %BOLD signal change between the Pretest and Posttest condition in the right superior temporal sulcus and the improvements in behavioral performance. Finally, a relationship between %BOLD signal change during training and the **difference** in %BOLD signal change between the Pretest and Posttest condition was observed for two frontal areas, the middle frontal gyrus and the inferior frontal gyrus.

Table 5. Correlation Results

Region of activation	% BOLD signal Change during training for all TRs correlated with difference in behavioral performance before and after training		Difference in % BOLD signal change between Pretest and Posttest for all TRs correlated with the difference in behavioral performance before and after training	
	p-value	**R-value**	**p-value**	**R-value**
Right Crus I	**0.01**	**-.7**		
Right superior temporal cortex			**0 .006**	**-0.7**

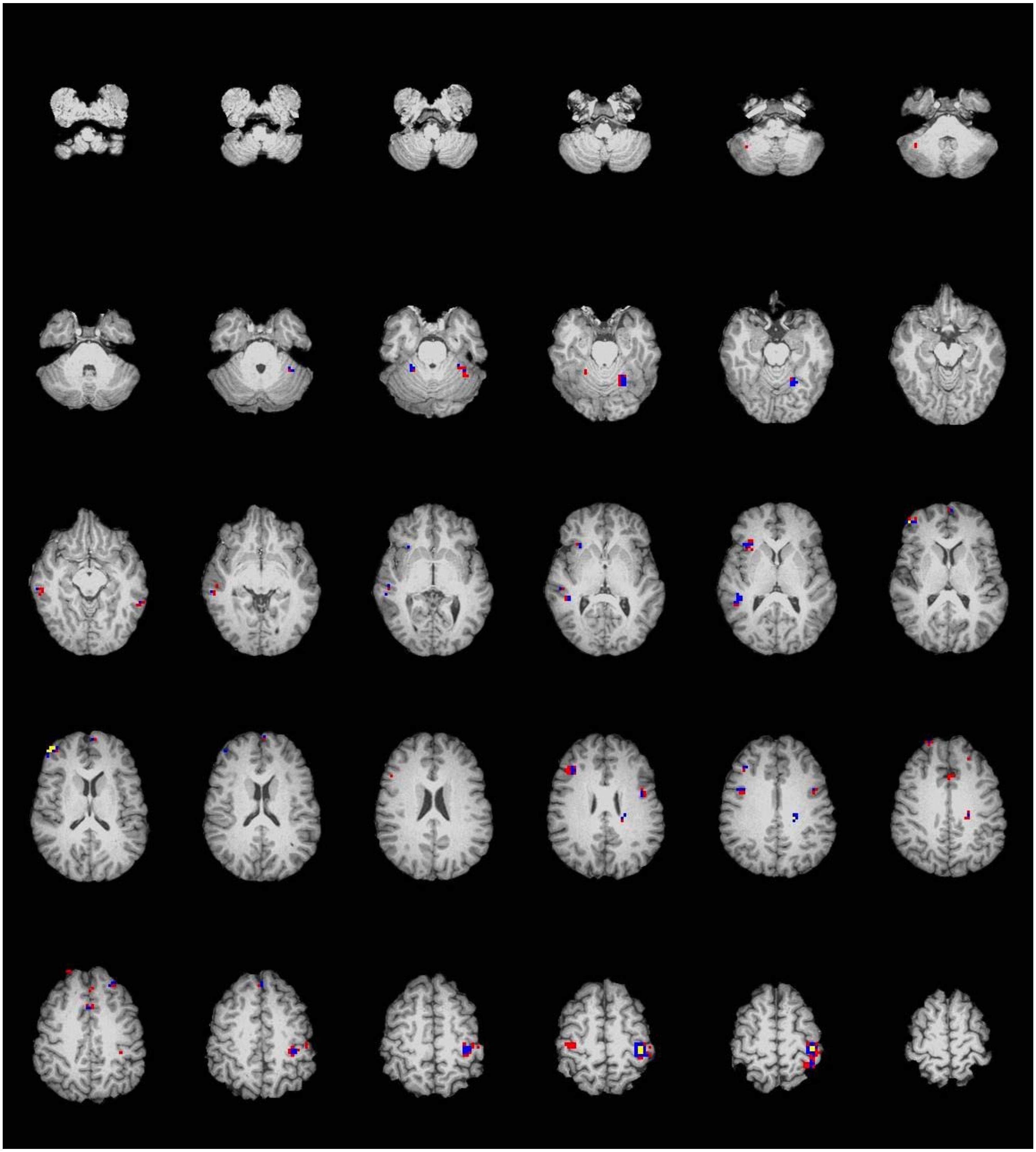


Figure 3. Pretest vs. Posttest Contrast

The axial montage displays regions of that were all significant active in a Condition (Pretest, Posttest) x TR. (1-11) interaction, p<0.001. Different colors reflect different levels of significance, all below 0.001.

Cerebellar Cortex. We identified four distinct regions in the cerebellum that showed a significant difference in brain activity when comparing before and after training. The right lateral hemisphere of the cerebellar cortex has been associated cognitive process including language (Fiez, 1996). Indeed one of the four regions we identified was located in right lateral Crus I (see Figure 4). This region responded early in the trial (see Figure 6a). In our correlation analysis, we found a significant correlation between the % change in BOLD response during training in this region with the difference on partial word accuracy score between pretest and posttest (see Table 5). Another region in right posterior Lobe V/VI (see Figure 5), showed a near-significant correlation, p=.06. The %BOLD signal change in this region occurred later and suggests that it is associated with the written motor response (see Figure 6b). A comparison between our written and no response trials confirmed this hypothesis (see Figure 6d). The percent change in BOLD responses for the remaining two left hemisphere cerebellar regions did not show a significant correlation with behavioral improvements in performance.

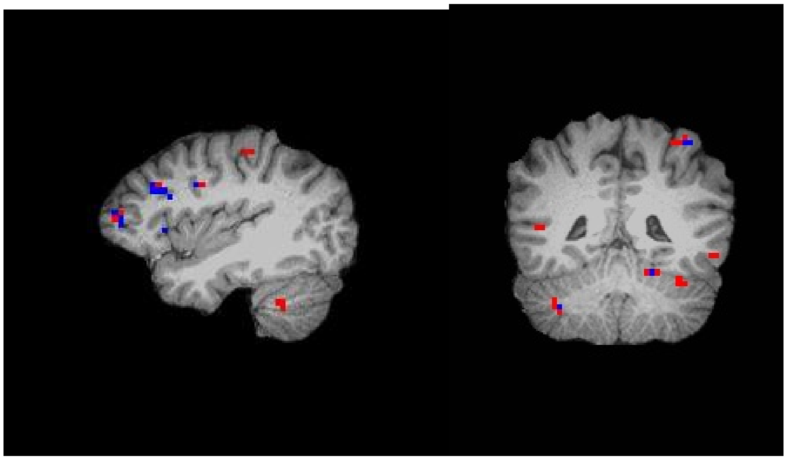


Figure 4. Right Crus I

Sagittal and Coronal view of right Crus I (left and right are reversed)

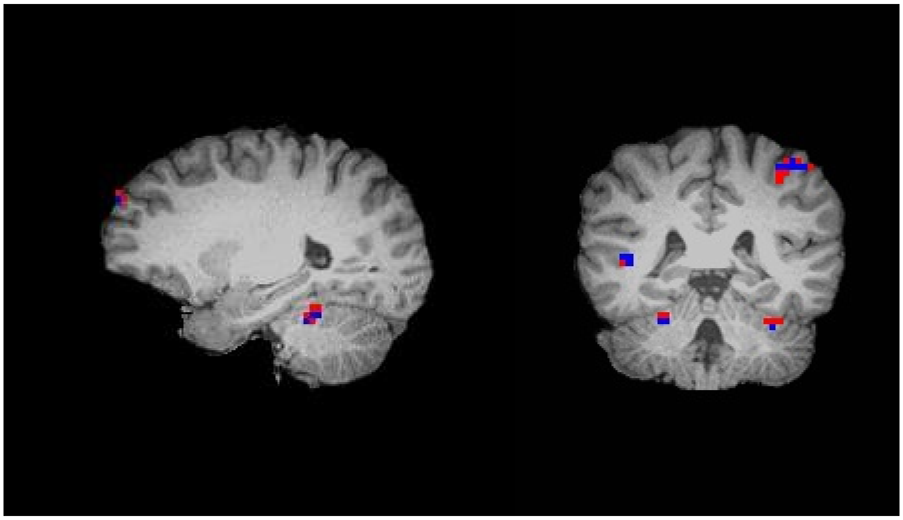


Figure 5. Right Lobe V/VI

Sagittal and Coronal view (left and right are reversed) of right Lobe V/VI

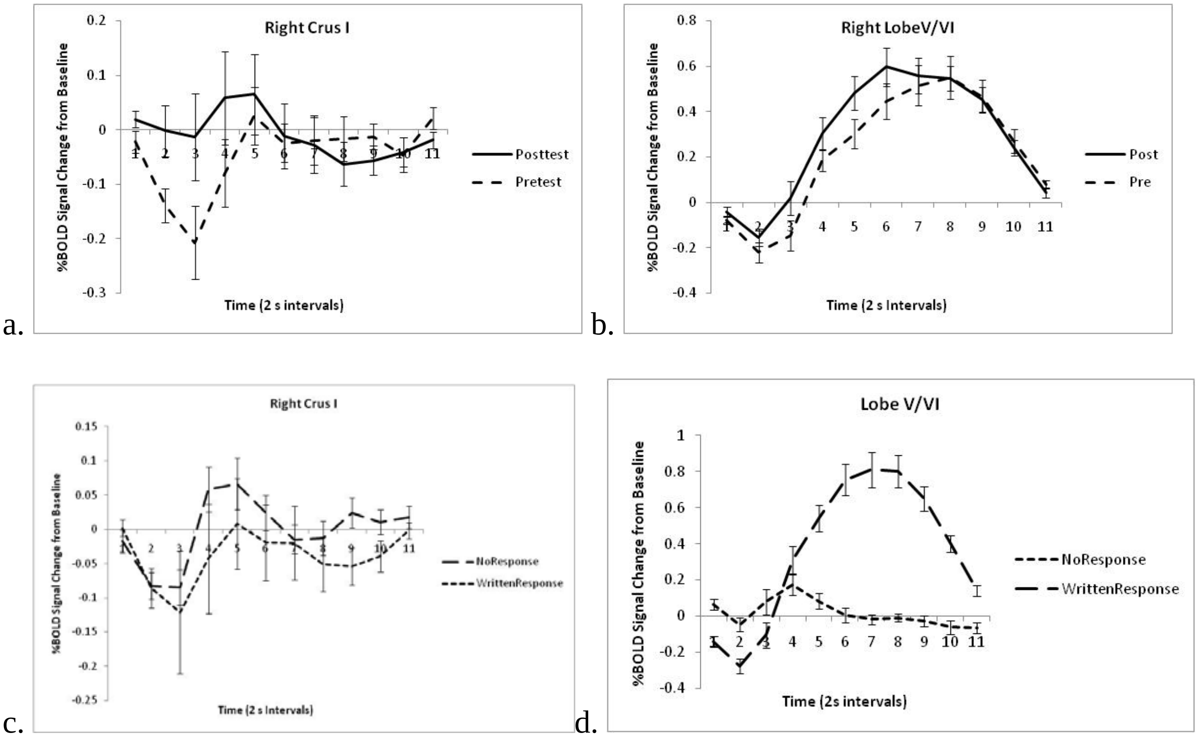


Figure 6. Timecourses of activation

a. Pretest vs. Posttest for Right Crus I **b.** Pretest vs. Posttest for Right Lobe V/VI **c.** Written Response vs. No Response for Right Crus I **d.** Written Response vs. No Response for Lobe V/VI

To further validate our findings for Crus I, we examined individual results for the same contrast. Displayed below are sagittal views for each participant that was included in the analysis (see Figure 7). In one subject, we were unable to determine the location of Crus I activation due to the shape of their cerebellar cortex. Three subjects did not show a cluster of activation that was within Crus I or II of the right hemisphere. However, the coronal view shows that activation clusters in these subjects was left-lateralized (see Figure 8). Individual variability in hemispheric lateralization for language and speech processing has been reported. Though often cerebral cortical activation is left-hemisphere dominant, the right hemisphere can be more dominant in some individuals (Fesl, et al., 2009).

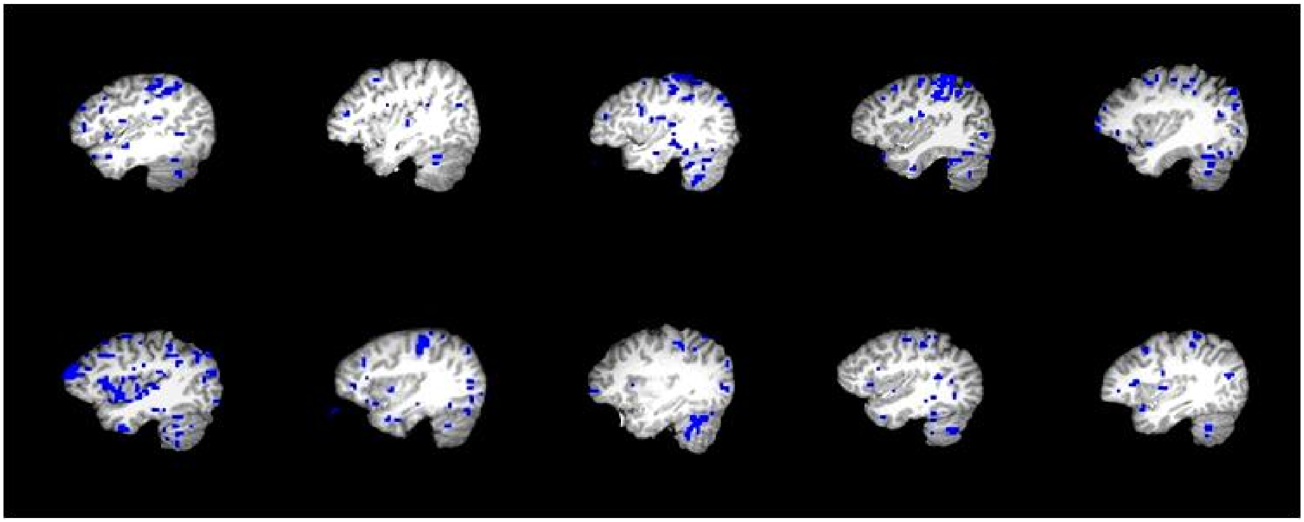
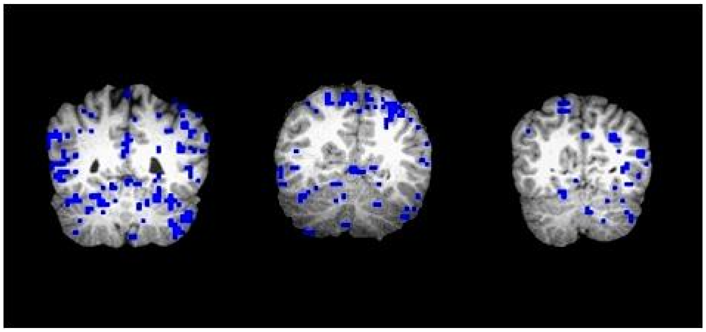


Figure 7. Individual Right Crus I activation

Sagittal view of Pre vs. Post contrast (condition x time interaction) in each individual subject (p<0.05). The majority of the subjects show a cluster of activation that falls in right Crus I or Crus II.

Figure 8. Pretest vs. Posttest Contrast

Coronal view of Pre vs. Post contrast (condition x time interaction) in three individual subject (p<0.05). Left and Right are reversed



Cerebral Cortex. We identified a number of regions in motor and premotor areas of cortex whose activation patterns were associated with the response period of the task (see Figure 3). We also found a regions in brain areas associated with executive control and attention such as the superior and middle frontal gyri, areas implicated in articulatory planning (inferior frontal gyrus and insula), as well as regions associated with linguistic representations (superior and inferior temporal cortex). Interestingly, one region in the posterior inferior/middle temporal cortex showed a near-significant correlation with changes in brain activity during training and improvements in behavioral performance. This region is located in the posterior part of inferior temporal cortex. In the Hickok and Poeppel model, inferior and middle temporal cortex is thought to be a component in the interface between lexico-semantic representations and sensory-based representations (Hickok & Poeppel, 2007). Therefore, it is not surprising that the activity in this region changes after training, as it may serve to guide the adaptation process.

We also found a region in the right superior temporal sulcus (see Figure 9). Changes in activity in this region between pretest and posttest significantly correlated with improvements in performance. In the Hickok and Poeppel model, this area is thought to be involved in the representation of prelexical phonological forms (sounds that map onto language-specific categories). Though their model suggests that both hemispheres contribute to this process, the fact that we see differences between pretest and posttest only in the right hemisphere and not the left is surprising since the left hemisphere is thought to be more involved in linguistic representations than the right hemisphere. This issue will be revisited in the General Discussion section.

Importantly changes in activity in this region between pretest and posttest significantly correlated with improvements in performance. The timecourse for this region is shown in Figure 10a as is the timecourse for the left inferior frontal gyrus Figure 10b, which has also been implicated in articulatory planning during speech production, motor imagery during speech perception, and attention. In Figures 11 and 12, we present example timecourses from other regions that showed a significant difference in our pretest and posttest contrast including motor and premotor cortex (see Figure 11) and frontal cortex (see Figure 12).

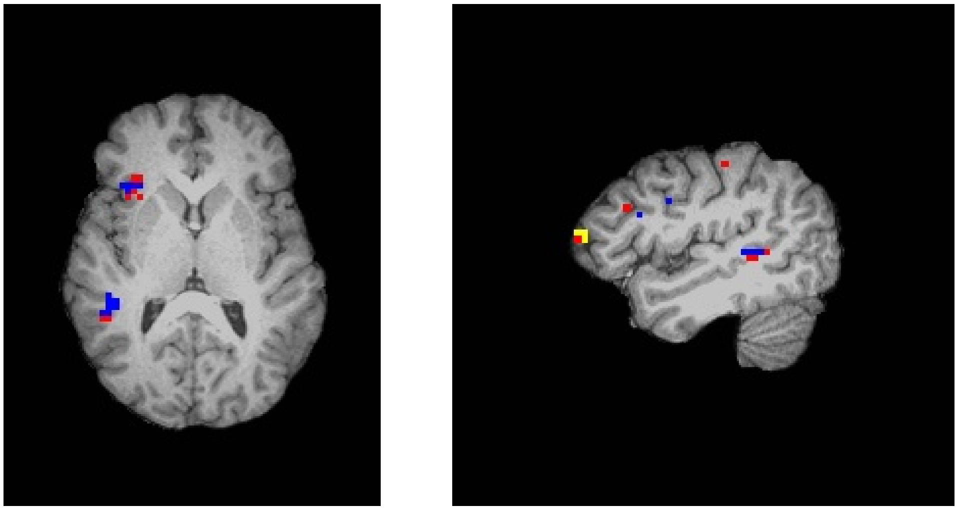


Figure 9. Right posterior superior temporal sulcus (STS)

Sagittal and axial view (Images are right and left reversed)

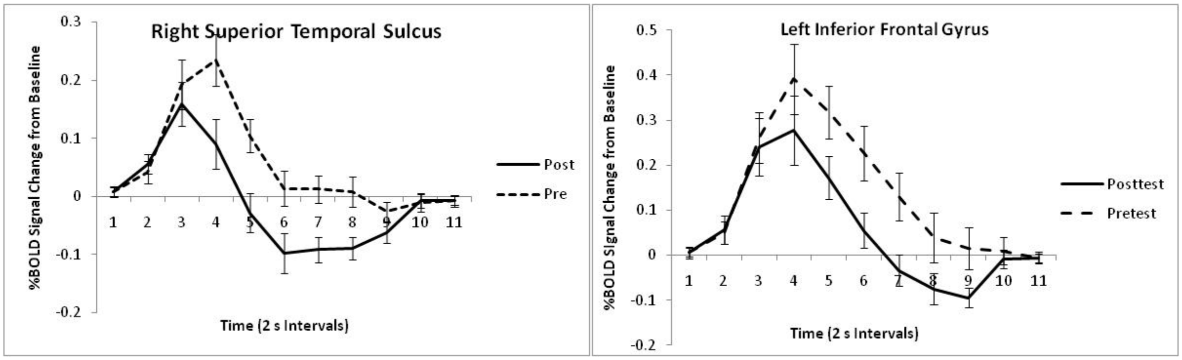


Figure 10. Example timecourses in temporal and inferior frontal cortex

a. Timecourse for the right superior temporal sulcus **b.** Timecourse for the left inferior frontal gyrus

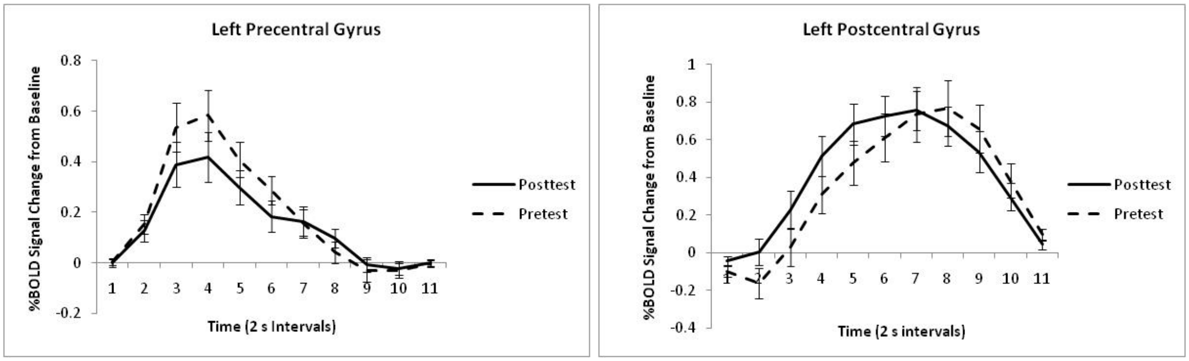


Figure 11. Example Timecourses in Motor and Premotor Cortex

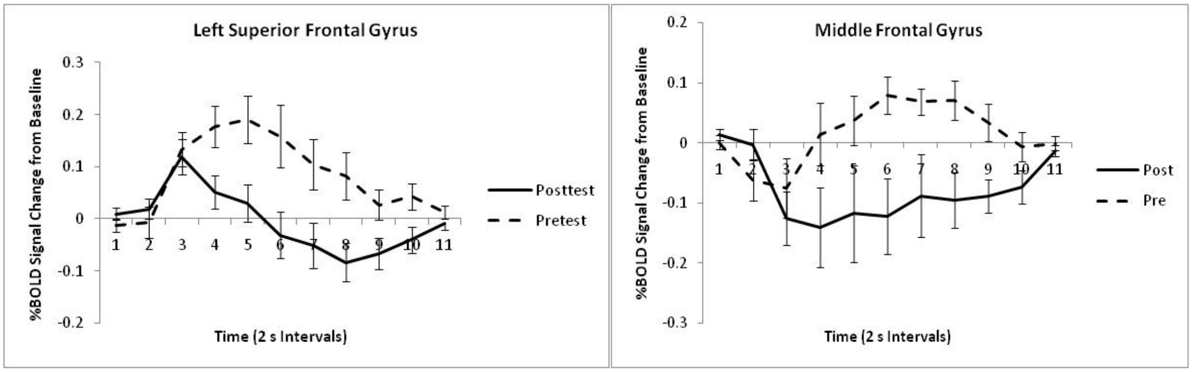


Figure 12. Example Timecourses in Frontal Cortex

3.3 DISCUSSION

Contributions of the cerebellum to sensorimotor adaptation have been a topic of interest that has sparked several theories of cerebellar function (Albus, Math Biosci; Thompson & Steinmetz, 2009; Braitenberg, 1967). Recently, theories have expanded to include functions beyond the motor domain (Ito, 2008; Doya, 2000). Though its involvement in cognitive tasks has been well-established, its specific involvement in a nomotor adaptation tasks has not been examined. Therefore, we examined behavioral and neural characteristics of an adaptive process in speech perception to determine whether we could find evidence for similar mechanisms of action when there is no intentional motor compensation. To avoid confounds related to timing mechanisms in the cerebellum, we chose a speech manipulation that only altered the spectral, but not the temporal components of the speech signal.

In general, auditory tasks are difficult to implement in the scanner due to unfavorable effects of scanner noise effects on auditory processing (Gaab, Gabrieli, & Glover, 2007; Zaehle, et al., 2007). Many different scanning procedures have been implemented to deal with issues related to scanner noise. However, many of these reduce statistical power and may lead to attentional confounds (Gaab, Gabrieli, & Glover, 2007). For this reason standard procedures were used in this study. Behavioral performance suffered as a consequence: errors were found during the perception of normal speech, and whole word recognition of the distorted stimuli remained very poor even after training. However, the fact that partial word accuracy improved following training indicates that access to lexico-semantic representations was sufficient to yield learning benefits.

A slow-event related design using standard imaging procedures enabled us to extract trial-level timecourses of activation from each of our regions of interest, making comparisons to behavior possible. To deal with attentional and scanner noise confounds, we focused our analyses on the contrast between the pretest and posttest conditions. Since these two conditions presented speech distorted at the same level of severity, changes in activity related to the acoustic properties of the speech signal would be eliminated by this contrast. Another advantage of our study is that we collected imaging data before training, during training, and after training. Therefore, we could establish significant differences between brain regions using pretest and posttest, and extract the activity of these regions during training to ascertain the relationship between brain activity during training and improvements in behavioral performance.

3.3.1 Localization and patterns of activity

Cerebellar Cortex

Advances in anatomical research on cerebellar organization and its anatomical connections to cortex emphasize the need for distinguishing between specific sites of activation within the cerebellum (Stoodley & Schmahmann, 2009; Strick, Dum, & Fiez, 2009). For example, cognitive, language, and speech perception tasks have all been associated with activation in the right lateral cerebellar hemispheres. Stoodley and Schmahmann compared cerebellar ROI's across a number of tasks and found that language-related tasks were localized to lobe VI and Crus I of the posterior lobe, whereas sensorimotor tasks tended to activate anterior lobe V and VI (Stoodley & Schmahmann, 2009). Regions identified in a meta-analysis of auditory tasks also identified a large cluster that extended from lobe VI to Crus I using an activation likelihood estimate (ALE) analysis (Petacchi, Laird, Fox, & Bower, 2005).

We identified a region in the right cerebellar hemisphere, in Crus I, that showed significant differences in BOLD signal before versus after training. In this region, we also found a significant correlation between the percent signal change of the BOLD response during training and improvements in behavioral performance. This was an especially interesting result, given the previous findings in the literature that indicate Crus I involvement in speech and language processes. We also identified a region in the right cerebellar hemisphere, in Lobe V/VI, that showed significant differences in BOLD signal before versus after training. In this region, correlation between the percent signal change of the BOLD response during training and improvements in behavioral performance did not reach significance.

New searchable databases (Research Imaging Center UTHSCSA, 2003-2009) allowed us to identify other studies that have reported similar foci of activation using Talairach coordinates as the basis for the search algorithm. We performed a search using the Talairach coordinates for the peak of the right hemispheric activation focus in Crus I and compared that to a search on the right Lobe V/VI peak locus (using a 7mm cubic search region). We found that prior contrasts involving a language manipulation yielded reported coordinates within the range of both regions; however, speech-specific studies only cited coordinates within the range of our Lobe V/VI region, and cognitive studies including working memory, spatial reasoning, and attention only cited coordinates that fell within the range of our Crus I region. This is consistent with our view that Crus I is involved in the adaptive process, whereas lobe V/VI is more involved in the motor output involved in the task.

These two regions also showed different activation patterns. The region in right lateral Crus I showed a % BOLD signal change early in the trial, which can be attributed to the stimulus presentation period. In this region, the % BOLD signal change is below baseline; the magnitude of the signal increases from pretest to posttest. The region in Lobe V/VI showed a % BOLD signal change that began later during the trial and extended through the response period. The BOLD signal change is this region was positive and also increased in magnitude from pretest to posttest.

Though the different patterns of activity in our cerebellar regions likely reflect different underlying neural mechanisms, the current literature does not provide us with a clear picture of the underlying cellular bases that cause changes in the BOLD response. The existing fMRI studies that report changes in cerebellar activity with motor learning and adaptation show increases, decreases, or do not report the direction in BOLD signal response (Jenmalm, Schmitz, Forssberg, & Ehrsson, 2006; Penhune & Doyon, 2005; Doyon, Song, Karni, Lalonde, Adams, & Ungerleider, 2002; Vaillancourt, Thulborn, & Corcos, 2003). Nonetheless, changes in blood oxygen level-dependent (BOLD) signal have often been used to infer changes in activity occurring in the cerebellum at the cellular level (Penhune & Doyon, 2005; Clower, Hoffman, Votaw, Faber, Woods, & Alexander, 1996; Diedrichsen, Hashambhoy, Rane, & Shadmehr, 2005).

The relationship between changes in BOLD signal and changes in neuronal activity is complex and is a subject of continued investigations (Logothetis & Wandell, Interpreting the BOLD signal, 2004). Simultaneous recordings of electrophysiological and fMRI data have shed light on their relationship (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). Local field potentials have been found to be the best predictors of changes in BOLD responses. Local field potentials are thought to represent an average of synaptic signals in a population of neurons, and may also include spike after-potentials and the voltage-gated membrane oscillations (Logothetis & Wandell, 2004).

Simultaneous measurements of neuronal activity, local field potentials, and cerebral blood flow can also be obtained during pharmacological manipulations. Thomsen et al. studied the effects of a GABAa blocker on neuronal and vascular signals (Thomsen, Offenhauser, & Lauritzen, 2004). The increases in Purkinje cell activity that resulted from the GABAa blocker were not associated with changes in cerebral blood flow. The authors concluded that cerebellar blood flow measures must reflect the excitatory synaptic activity from climbing fiber input and not Purkinje cell firing (Thomsen, Offenhauser, & Lauritzen, 2004). However, in another study, continued exposure to muscimol (an even more effective GABAa blocker) led to increases in local field potentials that could also be dissociated from changes in cerebral blood flow; the authors concluded that changes in cerebral blood flow may be more related to downstream processes caused by increases in calcium levels that follow excitatory input climbing fiber input

(Caesar, Thomsen, & Lauritzen, 2003). The effects of pharmacological manipulations on other cerebellar cell types and their resulting impact on cerebral blood flow, has yet to be studied. Thus our understanding of neural-vascular coupling within the cerebellum remains incomplete.

While it is not yet possible to attribute changes in the BOLD signal to specific neuronal components, the existing literature does point toward some speculations. If changes in BOLD signal are reflecting local field potential caused by excitatory synaptic inputs to Purkinje cells (such as those from climbing fibers), we might expect to see increases in the magnitude of the BOLD signal from the pretest condition (before training) to the posttest condition (after training), because there should be decreases in climbing fiber input associated with adaptation. However, our BOLD responses in our Crus I region of interest are negative, below baseline. Negative BOLD responses were examined in visual area V1 in monkeys and have been associated with decreases in neuronal activity and local field potentials (Shmuel, Augath, Oeltermann, & Logothetis, 2006). Thus, the increased magnitude in BOLD signal from pretest to posttest may reflect increases in neuronal activity. It is likely that participants continue to engage in an error correction process even after training since their accuracy scores are relatively poor. Therefore the changes in activity we observe may reflect other input to the cerebellum: for example, changes related to from mossy fibers in the cerebral cortex.

To summarize, our findings show evidence for right lateral cerebellar involvement in speech perception adaptation that is related to changes in behavior. One outstanding question that exists in the sensorimotor literature is whether specific cerebellar regions are recruited based on the types of representations involved or based on the task demands. For example, a lesion study in monkeys found that impairments in motor adaptation only occurred when lateral portions of the cerebellum were inactivated but not when motor areas in Lobe V/VI were inactivated (Baizer, Kralj-Hans, & Glickstein, 1999). Functional neuroimaging studies examining motor skill learning and adaptation, also report activation in right lateral Crus I (Doyon, Song, Karni, Lalonde, Adams, & Ungerleider, 2002).

Thus, it is possible that the recruitment of Crus I occurs for any adaptation regardless of the representation involved. This may reflect connections to areas in cortex that integrate different sources of information. On the other hand, even though Lobe V/VI and Crus I are often engaged in normal speech perception and language processing tasks (Petacchi, Laird, Fox, & Bower, 2005), our review of previously published studies suggested that distinctions can be drawn between these two regions, a finding that is consistent with the claim that lateral Crus I is involved more specifically in the adaptive processes engaged by our task.

Cerebral Cortex

In cerebral cortex, we found a network of regions that have previously been implicated in motor, language, and attention tasks. Thsee included regions in premotor and motor cortex, temporal cortex, and frontal cortex. We found one region in the posterior superior temporal sulcus that showed a significant negative correlation between changes in the BOLD signal and improvements in behavioral performance. Specifically the percent change of the BOLD signal in pretest vs. posttest contrast paralleled the improvements in prelexical phonological performance. We predicted that such a region should localize to areas previously implicated in phonological-level representations that are believed to closely interact with lexico-semantic representations, such as regions in superior temporal cortex. Indeed, we identified a region in superior temporal sulcus that showed a significant difference between the pretest and posttest conditions. We observed a decrease in hemodynamic response after training in superior temporal sulcus; the difference in %BOLD signal change correlated with improvements in behavioral performance. However, contrary to our predictions, this region was right-lateralized. Though the bilateral superior temporal cortices are throught to be involved in speech processing, the left hemisphere has been considered to be language-dominant (Hickok & Poeppel, 2007; Zatorre, Evans, Meyer, & Gjedde, 1992).

There are several possible explanations for the right hemisphere lateralization of the superior temporal sulcus region. The right hemisphere has been implicated in processing spectral information of the speech signal (Zatorre, Evans, Meyer, & Gjedde, 1992). Hence, our manipulation of the spectral parameters of the speech signal may have engaged the right hemisphere in the adaptive process due to the properties of the speech signal presented during training. Another possibility is that there are greater individual differences in the left hemisphere. There is precedent demonstrating that these individual variations in the left hemisphere temporal cortex can obscure group level effects (Okada & Hickok, 2006). Since our improvements in performance were variable, it is possible that our statistical power was insufficient in identifying regions in the left hemisphere.

We also identified a region in the posterior inferior/middle temporal gyrus in our Pretest vs. Posttest contrast. In the Hickok and Poeppel model, lexico-semantic processes are attributed to this region. Correlations between activity during training in the posterior inferior/middle temporal gyrus and improvements in behavioral performance in this region approached significance (p=.06). This is interesting, because our behavioral results in Chapter 2 suggest that accurate lexical access during training improves the adaptation process for spectrally shifted words and thus may help to guide the perceptual tuning.

Other regions identified in the Pretest vs. Posttest contrast included motor and premotor areas which likely demonstrate changes in the response production with adaptation. Finally, we found regions in frontal areas likely to be involved in attentional control mechanisms related to the adaptation process (e.g. regions in frontal cortex). Though attention caused by the speech distortion should be constant between the pretest and posttest condition, participants may be more differently engaged in the task before compared to after training. Their confidence about their responses and willingness to make a response could also change after training.

The correlations with behavior lead in the right superior temporal sulcus and in right Crus I lead to interesting predictions about the cerebro-cerebellar interactions that may underlie speech perception adaptation. Therefore, in the following chapter, we investigate these potential interactions between our regions of interest in the cerebellum, right Crus I and Lobe V/VI, and cerebral cortical areas through a functional connectivity analysis.

Acknowledgements

I would like to thank Kate Fissell for her help in optimizing the parameters for the imaging analysis and visualization, Jenna El-Wagaa for the phonetic coding of the behavioral responses, and Patryk Laurent for developing a program that we used to calculate the partial word accuracy scores. I would also like to thank Andreea Bostan and Dr. Richard Dum for sharing their expertise on cerebellar anatomy. Finally, I would like to acknowledge Scott Kurdilla and Deborah Vislay, the MR technications at the Brain Imaging Research Center, Corrine Durisko, our lab coordinator, and Charles Lotz, an undergraduate student, for their support.

4.0 FUNCTIONAL CONNECTIVITY ANALYSIS OF CEREBELLAR REGIONS INVOLVED IN SPEECH PERCEPTION ADAPTATION

In the previous chapter, we examined cerebellar involvement in speech perception adaptation and identified cerebellar regions of interest using a pretest to posttest contrast. A region in the right lateral hemisphere, in Crus I, showed individual differences correlations in brain activity during training that were related to individual differences in behavioral improvements before and after training. Another region in the right cerebellar hemisphere, in Lobe V/VI, was associated with the response period of the task. The different response properties of these two regions lead us to propose that they are part of two distinct cerebrocerebellar networks. If this is the case, then the two cerebellar regions should have different patterns of neuroanatomical connectivity with cerebral cortex. In this chapter, functional connectivity analyses are used to address this question.

The anatomical connections between the cerebellum and motor cortex are well-characterized. Closed cerebro-cerebellar loops project from Lobe V/VI to Motor cortex and travel through the dorsal dentate and the ventrolateral thalamus to form the basis of the interactions between motor areas in the cerebellar cortex and motor cortex (Dum & Strick, 2003; Kelly & Strick, 2003). Recently established cerebro-cerebellar connections between the cerebellum and other cortical regions suggest a potential anatomical basis for cerebellar involvement in non-motor processes. Specifically, recurrent loops between prefrontal cortex and the lateral cerebellum have also been anatomically defined (Kelly & Strick, 2003).

Advances in functional neuroimaging methods have contributed to this literature by using functional connectivity analyses to explore connections between the cerebellum and cerebral cortex (Fox & Raichle, 2007). Simple functional connectivity analyses explore correlations between activity in one region and activity across the rest of the brain. They have also been used to examine correlations between two pre-determined regions of interest. While there are many caveats to these types of analyses, they offer a non-invasive method for examining potential anatomical connectivity in the human brain. One approach is to use functional imaging data collected during a “resting state”. This approach has been used in several studies to explore cerebellar connections to cerebral cortex (Allen, McColl, Barnard, Ringe, Fleckenstein, & Cullum, 2005; Krienen & Buckner, 2009; O'Reilly, Beckman, Tomassini, Ramnani, & Johansen-Berg, 2009). The advantage of this method is that perturbations associated with task-specific changes in connectivity or correlations do not contribute to the correlations observed between the regions. Thus, this approach has been touted as a way to determine the “intrinsic connectivity” between brain regions.

Another approach is to probe for connectivity using data collected during the performance of an active task. For example, one study examined connections between three brain regions (including a region the cerebellum) during a rhyme judgment task (Booth, Wood, Lu, Houk, & Bitan, 2006). One advantage of this approach is that the regions of interest are defined by the task, and thus there is a strong *a priori* reason to be interested in their functional connectivity. We used this second approach to explore connections between two seed regions in the cerebellum: the right lateral Crus I and the Lobe V/VI regions identified in Chapter 3. The former was chosen because it exhibited fluctuations in the BOLD response during the early portion of each trial and because significant correlations between brain activity during training, and improvements in behavior were found in this region. The latter region was chosen because it exhibited a motor-related response, and based upon its locus within the cerebellum, functional connectivity to motor cortex would be expected. Thus, the goal was to use this region to help validate the results for the Crus I region. We used a simple functional connectivity analysis to probe for potential functional connections between our two seed regions in the cerebellum and voxels located throughout the rest of the brain. The functional images obtained during our training condition provided the data for the connectivity analysis. This is a further strength of the study, because the connectivity analysis was thus performed using data that were independent from the Pretest vs. Posttest data used to identify our seed regions.

4.1 METHODS

We used our Crus I and Lobe V/VI voxel clusters from the pre vs. post training contrast as seed regions for our functional connectivity analysis. The data were taken from the 14 participants analyzed in the functional neuroimaging study in Chapter 3. Average timecourses (mean BOLD signal intensity for each trial and at each timepoint) were extracted from the voxels that comprised the seed regions; these timecourses came from images that were corrected for linear drift, motion-corrected, and smoothed to a common reference space during the preprocessing analysis in Chapter 3. Each subject's timecourse for each seed region was correlated with the timecourse for each voxel in the smoothed and corrected images. The resulting coefficient values for each voxel from images for each participant were subsequently used as dependent measures for voxelwise t-tests that assessed the degree to which the coefficient values differed from zero across participants. The resulting T-map was visualized using AFNI at a p-value < 0.01 with a corrected cluster size that was determine by using the AlphaSim program in AFNI (Ward, 2000).

4.2 RESULTS

Our simple functional connectivity analysis examined correlations between timecourses extracted from two seed regions: right Crus I and right Lobe V/VI that were defined in Chapter 3. The regions that were correlated with Crus I are reported in Table 6 and shown in Figure 13. The regions that were functionally connected to the Lobe V/VI seed region are reported in Table 7 and shown in Figure 14. The right lateral Crus I seed showed significant functional connections with a small number of regions including the left superior temporal plane and the left angular gyrus.

Our results demonstrate that the most significant functional connections to the right Lobe V/VI seed region include motor and somatosensory cortex. These results contrast with the functional connectivity map obtained with the right lateral Crus I seed region, which shows no significant connectivity to motor cortex even at higher p-values. The other regions in the connectivity map for right Lobe V/VI included supplementary motor area, inferior parietal, and superior frontal cortex.

4.2.1 Right Crus I

Table 6. Correlations with right Crus I

Regions that are correlated with right Crus I, corrected p< 0.01

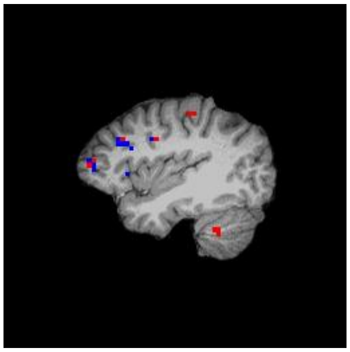
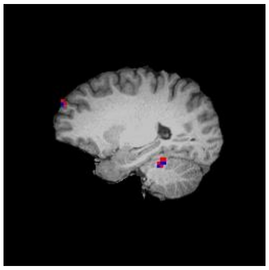
Seed Region		Talairach Coordinates	Size
Crus I (Image Taken from Chapter 3) 	Right Crus I extending all the way to the Left Crus I	36, -48, -38	3492
	Right Hippocampus	29, -25, -4	314
	Angular Gyrus/BA39	-37, -57, 28	124
	Left Spt and Heschl’s gyrus	-51, -21, 18/-38, -24, 18	107

Table 7. Correlations with right Lobe V/VI

Regions that are correlated with right Lobe V/VI, corrected p< 0.01

Seed Region		Talairach Coordinates	Size
LOBE V/VI (Image Taken from Chapter 3) 	Superior Frontal Gyrus/BA9	-13, 54, 35	5293
	Right cerebellum V/VI	22, -38, -22	3468
	Left inferior parietal, postcentral gyrus, precentral gyrus	-33, -43, 55	778
	Left postcentral gyrus	33, -29, 45	747
	Right putamen	22, -4, 10	
	SMA/BA6	0, -15, 51	159

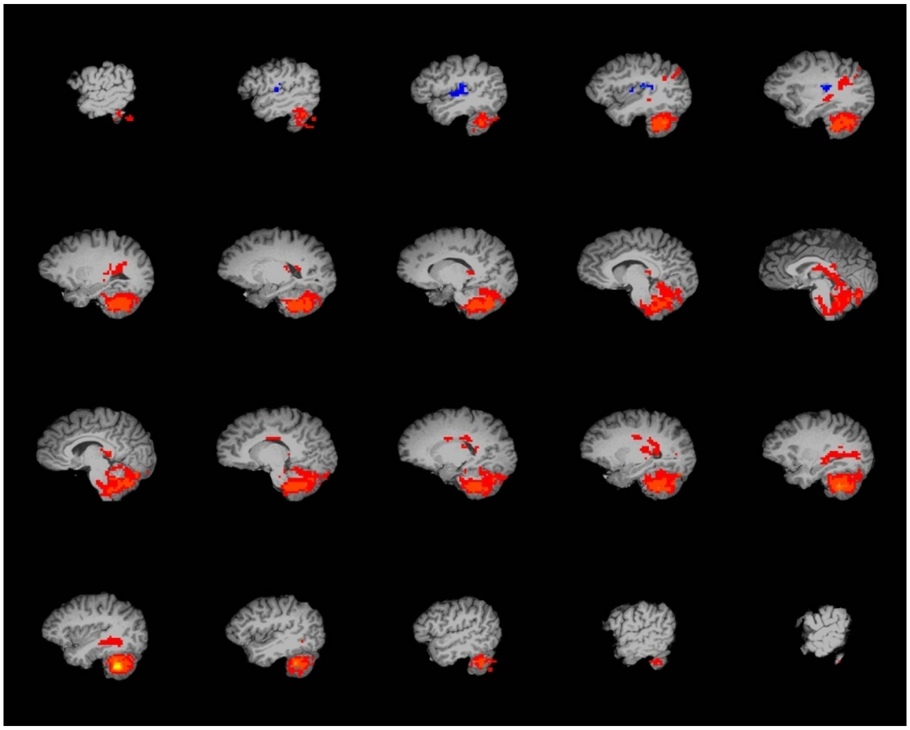


Figure 13. Functional Connectivity for Right Crus I

This figure shows significant correlations at a corrected p-value <0.01

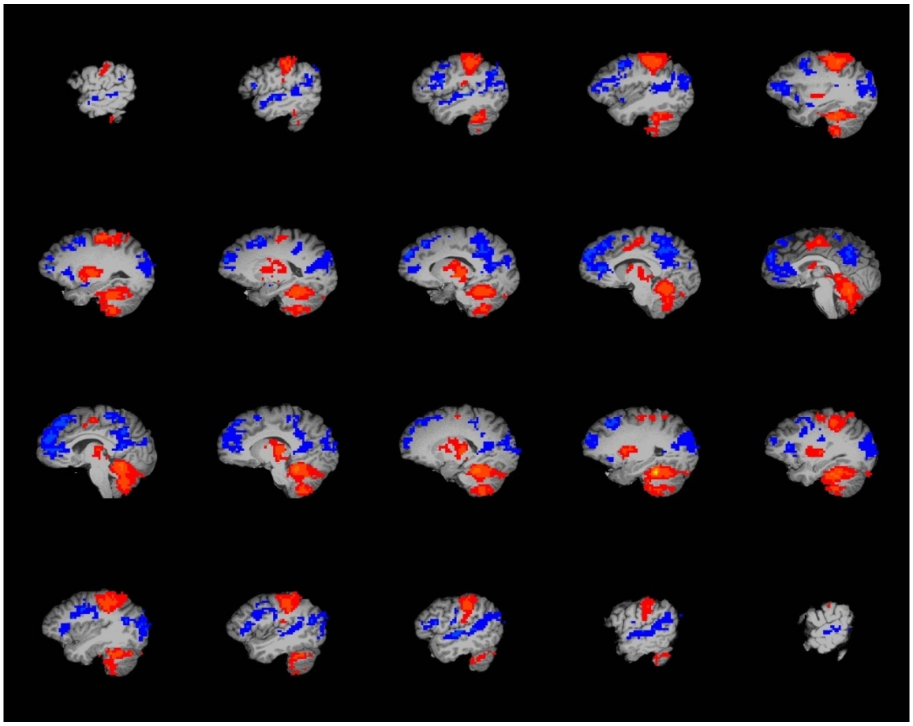


Figure 14. Functional Connectivity for Right Lobe V/VI

This figure shows significant correlations at a corrected p-value <0.01 corrected

4.3 DISCUSSION

In Chapter 3 we identified a specific region in the cerebellum that was involved in a speech perception adaptation task -- a right lateral region in Crus I. We speculated that it may be engaged in an error correction processes that is analogous to that described for sensorimotor adaptation. To gain insight into the potential interactions between the cerebellum and cerebral cortex that might drive speech perception adaptation, we performed a simple functional connectivity analysis. Based on our results in cerebral cortex from Chapter 3, we speculated that regions associated with prelexical phonological (e.g., area STS) or lexico-semantic representation (e.g., the middle temporal gyrus) s would be good candidate regions for functional connectivity to Crus I. In contrast, we predicted that a region in Lobe V/VI, which was associated with the motor response, would show evidence of functional connectivity with regions involved motor execution.

We found functional connections between our Crus I seed region and area Spt, but contrary to our predictions, we found no evidence of functional connectivity to area STS or other regions within the middle and inferior temporal gyrus. A further complication is that our Pretest vs. Posttest contrast in Chapter 3 did not reveal significant differences in activity for area Spt. However, inspection of our simple main effect of time images from our Training condition does indicate that area Spt is active during the perception of distorted speech input. In the final chapter, alternative frameworks that may be able to account for these seemingly discrepant findings are discussed.

Previous studies using resting state functional connectivity analysis used cerebral cortical regions as seed regions and revealed that Lobe V and VI in the cerebellum are functionally connected to motor, somatosensory, auditory, and visual areas, whereas Lobules VIIa, Crus I, and II are functionally connected to prefrontal and posterior-parietal cortex (O'Reilly, Beckman, Tomassini, Ramnani, & Johansen-Berg, 2009). Based on their results, O'Reilly et al. propose a distinction between sensory and motor projections to the cerebral cortex, which may be overlapping, and separate connections between non-motor representations and cerebral cortex, which they refer to as coming from a "supramodal zone". Another study specifically explored connections in Crus I and Crus II to predefined regions within prefrontal cortex; this study found differences in the strengths of the correlations that depended on the location of the seed region (Krienen & Buckner, 2009). Though we replicated the basic finding of Lobe V/VI connectivity to motor cortex, we did not find connections between our seed region in Crus I and prefrontal cortex. There are two major differences between the two previous studies that may account for this discrepancy. The first possibility is that our region was a specific region within Crus I that was defined by a specific task contrast, whereas in the Krienan study, the seed region was based on a connectivity map derived from prefrontal cortex. As noted in their study, functional connectivity analyses may extract regions that are indirectly activated by other regions in cerebral cortex. Viral tracing studies have shown that projections to different non-motor areas of the cerebral cortex are topographically organized in the ventral dentate, so it is possible that nearby regions within Crus I could have different patterns of cerebral connectivity (Dum & Strick, 2003; Middleton & Strick, 1998).

Our goal for this study was to identify potential pathways that may facilitate speech perception adaptation. Thus for our purposes, either indirect or direct connections from the cerebellum to cerebral cortex would be interesting. We found differential patterns of activity for our two distinct seed regions. The region in Crus I was functionally connected to a region that seemed to encompass both Hesch's gyrus and a portion of the superior temporal plane that is thought to be involved in sensorimotor speech processing (area Spt, as denoted by Hickok and Poeppel's 2007 model). We found no significant functional connectivity between Crus I and motor cortex. In contrast, the region in Lobe V/VI did show functional connections with motor cortex. These results are consistent with the emerging view that motor and non-motor representations are anatomically segregated within the cerebellum (Dum & Strick, 2003; Kelly & Strick, 2003).

The results for Lobe V/VI replicate the existing literature, which demonstrates "intrinsic" functional connections between Lobe V/VI and sensory and motor areas in cerebral cortex obtained through cerebral cortical seed regions. The ability to replicate these results gives some credence to the methods we used in our study and make our findings for Crus I more convincing. We also note that anatomical tracing studies in non-human primates have provided some anatomical evidence for connections between area Spt and the cerebellum (Schmahmann & Pandya, 1991).

However, there are limitations to our methodology; it has been argued that correlation analyses using resting state data are better indicators of anatomical connectivity, because the correlations are more likely to reflect the intrinsic, task-independent, properties of the neuronal system (Fox and Raichle 2007). Since our functional imaging data were obtained within the context of a task, it is possible that similar trial-level fluctuations in the BOLD signal (e.g., phasic increases related to stimulus presentation) could account for much of the co-variation. Thus, correlations obtained during the context of a task could reflect regions simultaneously activated by the task even if they are not part of the same intrinsic functional network.

On the other hand, an advantage of our approach is that the seed regions were specifically selected on the basis of data that implicated this region is speech perception adaptation (for the

Crus I region) or motor output (for the Lobe V/VII region). More specifically, the Crus I seed region was defined by a specific task contrast that examined free word recognition on a severely distorted speech condition, before and after training. This region also showed a significant correlation with improvements in behavioral performance, which led us to propose that it may contribute to an error correction process in speech perception adaptation. Thus, the functional connections observed between Crus I and Spt, an area already known to be involved in speech processing is a particularly exciting finding that can lead to future testable hypotheses. The results suggest that Spt may participate in an error correction process, either through a direct or indirect pathway that engages the cerebellum. The functional connectivity to angular gyrus is more surprising since this region falls outside the domain of the speech processing network reviewed by Hickok and Poeppel. However, it is known to be involved in language processing, and classic models of language processing have long held that it serves as an interface between written language and spoken speech (Carreiras, et al., 2009). Thus, the potential involvement of this region in speech perception adaptation may fit within its broader contributions to language processing, or may be indirectly correlated through another area. In the following chapter, we discuss in greater detail how our functional connectivity results, in combination with our behavioral and functional activation results, might be brought together to support alternative accounts for how the cerebellum may support speech perception adaptation.

Acknowledgements

I would like to thank Sung-Joo Lim and Patryk Laurent for contributing to the functional connectivity analysis.

5.0 GENERAL DISCUSSION

Functional neuroimaging has significantly advanced our knowledge of cerebellar contributions to cognition with numerous studies reporting changes in activity related to learning, language, memory, and attention (Stoodley & Schmahmann, 2009; Strick, Dum, & Fiez, 2009). Despite the accumulating evidence supporting the role of the cerebellum in cognition, the nature of its involvement in these tasks remains unclear, and is often attributed to a motor planning component (Desmond, Gabrieli, Wagner, Ginier, & Glover, 1997; Glickstein, 2007). In contrast, a more comprehensive understanding of cerebellar involvement in sensorimotor adaptation has led to sophisticated theoretical models of cerebellar function, which can, in theory, be applied to non-motor tasks (Ito, 2008). An empirical examination of an adaptive process that is guided by non-motor processes could help to bridge the gap between these two areas of research. Towards this end, we used functional magnetic neuroimaging study to investigate speech perception adaptation. Speech perception adapts to distortions in the speech input, and according to theoretical models, the adaptive process is guided by cognitive information about word identity (lexico-semantic information) (McClelland, Mirman, & Holt, 2006; Norris, McQueen, & Cutler, 2003).

5.1 IS THE CEREBELLUM INVOLVED IN SPEECH PERCEPTION ADAPTATION?

In a free-word recognition task, we examined changes in brain activity before, during, and after training on spectrally distorted words. This manipulation allowed us to control for cerebellar activity that was related to differences in time perception may also engage the cerebellum (Ackermann, Graber, Hertrich, & Daum, 1997). To specifically examine effects related to the adaptation process, several factors were taken into consideration. For example, differences in activity could be due to the distorted acoustic input or differences in cognitive aspects of the task reflecting increased difficulty in performing the task. A functional neuroimaging study examining adaptation during visually-guided reaching by Clower et al. emphasized the importance of accounting for activity related to the error detection/correction and not the adaptation (Clower, Hoffman, Votaw, Faber, Woods, & Alexander, 1996). Therefore, we focused our analysis on regions that exhibit a significant pretest vs. posttest difference in BOLD signal, based upon the rationale that sensory and cognitive components of the task that are unrelated to the adaptation process would be accounted for by holding constant the type of stimulus that was delivered. As a secondary condition, we required a link between changes in brain activity and changes in behavioral performance. These two conditions were met for a region in the right lateral Crus I area. Specifically, this region showed significant changes in BOLD signal in a pretest to posttest contrast. Moreover, changes in activity in this region during training correlated with the behavioral improvements in performance observed when comparing performance that followed training (*for discussion see* Chapter 3). Taken together, these results support a behaviorally-relevant role for this region of the cerebellum in speech perception adaptation.

5.2 SPEECH PERCEPTION ADAPTATION

5.2.1 Error Signals

Our results suggest that the cerebellum is involved in speech perception adaptation. To understand how the cerebellum may "retune" perceptual processing, it is useful to consider what has been learned in the sensorimotor domain. Theoretical accounts of cerebellar contributions to sensorimotor adaptation have emerged from a considerable amount of evidence. Notable areas of investigation include studies of eye movement control, visually-guided reaching, and classical conditioning. Several leading accounts are based on a common underlying theme that the cerebellar circuitry is involved in prediction error and error correction processes (Wolpert, Miall, & Kawato, 1998; Doya, 2000; Albus, Math Biosci; Schultz & Dickinson, 2000). Computationally-based accounts implement this process through internal models that mimic the dynamics of the sensory and motor representations (Wolpert, Miall, & Kawato, 1998).

In the sensorimotor literature, internal models consist of forward models that predict sensory consequences based on a motor command, and inverse models that predict the motor command needed for the desired sensory outcome. In both cases, the predicted state is compared to the actual state. The resulting discrepancies generate internal error signals, which can be used to retune the mapping between sensory and motor representations (Wolpert, Miall, & Kawato, 1998). A translation of this basic process was explored by Ito for the case of mental thought (Ito, 2008). In his instantiation of the model for cognitive processes, either a forward or an inverse model controls the dynamics of the controlled object (in this case a mental model rather than a physical body part); the mental model could be generated from "perception, imagination, or the comprehension of discourse" (Ito, 2008). Similarly, discrepancies between the intended and actual outcome of the mental model can generate error signals which can tune these dynamics.

The details of the remapping process between sensory and motor representations in the sensorimotor domain are still under debate. The inferior olive is often assumed to be the site of comparison since it is thought to provide the error signals to the cerebellum through climbing fiber input (Bloedel & Bracha, 1998). Consistent with this view, electrophysiological recordings have shown that complex spike activity signals the direction and size of saccade errors (Soetedjo, Kojima, & Fuchs, 2008); thus, successful adaptation can be reflected by a decrease in climbing fiber input. Recent studies examining short-term saccadic adaptation have also demonstrated that changes in the Purkinje cell simple-spike population responses reflect the direction of the gain, and suggest that adaptation results from "optimizing the shape" of the Purkinje cell simple spike population response (Catz, Dicke, & Thier, 2008). However, there are still many unknowns including anatomical connections between cerebral cortex and the inferior olive, and the representational space of the error code (Kawato & Gomi, 1992; Wolpert, Miall, & Kawato, 1998).

With so many unknowns in the sensorimotor domain, any conclusions about the mechanisms involved in speech perception adaptation must necessarily be highly speculative. Internal models for speech perception may be represented either within the cerebellum or through interactions between cerebral cortical areas that convey internal models to the cerebellum or the inferior olive. Despite these limitations, progress in generating testable hypotheses can come from examining the dynamics of the entire cerebro-cerebellar system. Thus, we propose several potential pathways that could facilitate representational comparisons that could in turn guide error correction in speech perception adaptation.

For the purposes of our discussion, we use Hickok and Poeppel's 2007 model of speech processing as our framework for mapping speech representations onto corresponding brain structures. While this model has gained a great deal of popularity in the language domain, it is important to note that certain components of the model remain speculative in nature whereas others continue to be supported by an increasing number of studies on phonological processing (Desai, Liebenthal, Waldron, & Binder, 2008; Leech, Holt, Devlin, & Dick, 2009; Okada & Hickok, 2006). In the Hickok and Poeppel model, following spectro-temporal analysis of the acoustic input, speech accesses a phonological network that diverges into two streams: a ventral stream involved in accessing representations for comprehension, and a dorsal stream involved in accessing representations for articulation. These two streams lay the foundation for a system that can engage alternative pathways in order to make comparisons between actual and desired outcomes of speech processing. We consider two possible pathways that could form the basis of such comparisons.

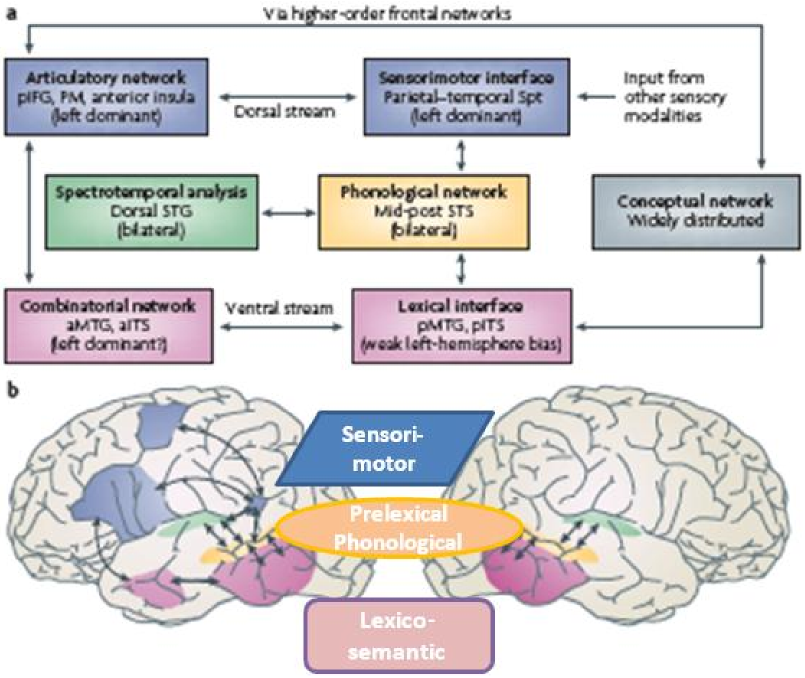


Figure 15. Modified Speech Processing Model

Source: Reprinted by permission from Macmillan Publishers Ltd Nature Neuroscience (Hickok and Poeppel, 2007). This figure was modified to illustrate the relationship with the terminology and processes we propose for speech perception adaptation.

In the first proposed pathway, the distorted incoming sensory information activates multiple and competing prelexical phonological representations in STS. After the participant selects a word, the chosen lexico-semantic representation of the word in turn activates an unambiguous set of prelexical phonological representations through interactive processing between STS and regions associated with lexico-semantic representations in the middle and inferior temporal cortex. The selected prelexical phonological representations can in turn predict the sensory outcome through an internal model, and be compared with the original sensory input. To resolve the error, the discrepancy in sensory information would alter the internal model's predictions about the relationship between prelexical phonological and sensory information.

The mechanisms that are proposed for this pathway can be used to explain the behavioral correlations that we found in Chapter 3. Our functional neuroimaging results provided evidence of a near-significant correlation between improvements in behavior and the magnitude of the BOLD response in the left posterior inferior/middle temporal gyrus region during training. If pITG/MTG provides the lexical information that guides adaptation, then it would make sense that greater activation in this region could yield a larger training benefit (i.e., larger difference in PreTest vs. PostTest activation). A similar correlation between the level of activity during training and subsequent improvements in speech perception performance was also found in right Crus I. This finding could be taken as evidence that Crus I is involved in the error-correction associated with adaptation: i.e., increased error correction during training could lead to increased adaptation and greater pre vs. post training performance benefits. A third significant correlation with behavior was found in the right superior temporal sulcus, but in this case the correlation was between improvements in behavioral performance and changes in brain activity in PreTest vs. PostTest contrasts. This finding could be consistent with the idea that STS is the site of adaptation. As the prelexical phonological representations become retuned, the acoustic input may be able to provide more efficient access, resulting in a reduction in synaptic activity and a decrease in the overall BOLD response.

Our results also point to two potential problems with this proposed pathway. The first problem is that both the Crus I region and the STS region we identified in our behavioral correlations were in the right hemisphere. This is inconsistent with the cross-hemispheric anatomical connections that underlie cerebro-cerebellar interactions. The second problem is that our functional connectivity results for the Crus I seed region do not show functional connections to STS. Instead, we found significant functional connections between right Crus I and area Spt in the left hemisphere, an area likely to involve sensorimotor representations. Therefore, we consider an alternative pathway for speech perception adaptation that is mediated through sensorimotor representations.

An alternative scheme considers how the adaptive processes could instead be driven by sensorimotor processes. In this example, the actual distorted sensory input is transformed into approximate sensorimotor representations in area Spt. Concurrently, lexical representations generated from the perceived word could also be used to access non-distorted sensorimotor or articulatory-based representations of the perceived word via the ventral pathway of the Hickok model (see Figure 15). These motor-based representations could in turn be used to predict the sensory (or sensorimotor) consequences of the motor commands. Finally, the predicted representations could be compared to the actual sensorimotor representations generated within Spt in response to the acoustic input. The fact that area Spt does not show significant changes in activity between a pretest and a posttest condition may seem puzzling. However, it is possible that remapping does not necessarily alter overall levels of synaptic activity, in which case an overall change in the BOLD response would not be expected. Areas that are interconnected with Spt (e.g., STS) could potentially benefit from the remapping that has occurred; for example, sensorimotor remapping in Spt could increase the accuracy and efficiency of prelexical phonological representation within STS, leading to changes in the BOLD signal within STS that correlate with sublexical changes in perceptual accuracy. As in the first pathway, the learning would be guided by successful lexical access, though in this case the lexico-semantic representations serve as a bridge into the motor system rather than as the direct source of a mental model for error correction. Thus, as in the first pathway, the near-significant correlations between activity in inferior temporal cortex during training and perceptual learning gains could indicate that more successful lexico-semantic access provides better guidance for learning, and a similar relationship in Crus I could indicate that this increased guidance leads to more effective error correction by the cerebellum. While the error correction in the first pathway was directed to STS, in this second pathway the error correction would be directed to Spt. Changes in Spt would in turn facilitate the prelexical phonological processing in area STS.

Similar ideas have been suggested in verbal working memory, where the to-be maintained information is rehearsed through articulatory representations and the “trajectories” (predicted sensory consequence) of those articulatory representations are compared to the maintained representations in temporoparietal cortex (Desmond, Gabrieli, Wagner, Ginier, & Glover, 1997). In the speech production literature, Guenther et al. also proposed a model where predicted sensory consequences of motor commands are used to correct motor output. Though the cerebellum is included in their model, its exact role in the feedforward and feedback control systems is not well-defined (Guenther & Ghosh, 2003)

In summary, both adaptation pathways are guided by lexico-semantic representations which inform the prelexical phonological choice. Throughout adaptation progresses, the prelexical phonological choice tunes acoustic-prelexical phonological mappings (pathway 1), or acoustic-sensorimotor mappings (pathway 2). Successful adaptation leads to a more accurate prelexical phonological choice. Thus, more intelligible stimuli should lead to better the lexico-semantic guidance, and more successful error correction and adaptation. With iteration, adaptation should eventually lead to improved perceptual accuracy at the whole word level, unless the initial distortion is so severe that it causes insufficient or extremely noisy access to lexico-semantic representations.

5.2.2 Anatomical Pathways

Our previous discussion suggested two possibilities for the representations involved in generating the error: acoustic-prelexical phonological and acoustic-sensorimotor. However, how the error information is propagated to specific anatomical structures and used to retune these relationships is unclear. The inferior olive has been suggested as one possibility for the site of comparison between sensory and motor representations. However, others have suggested that the error can be represented in cerebral cortex (Guenther & Ghosh, 2003). We propose three possible anatomical bases for interactions between the cerebellum and cerebral cortex. We consider these three possibilities for the second pathway we proposed above where the perceptual retuning is mediated through area Spt.

Since cerebro-cerebellar interactions have been shown to form closed loops, we preserve this structure in considering three alternative anatomical pathways recuirted by speech perception adaptation. In the first one closed cerebrocerebellar loop mediates this process. Through anatomical tracing studies in non-human primates, Schmahmann et al. has demonstrated that the temporoparietal junction and the upper bank of the superior temporal sulcus project to the pontine nuclei (the major source of input to the cerebellum) (Schmahmann & Pandya, 1991). Thus, our first model suggests that the cerebellum projects directly to area Spt to retune acoustic-sensorimotor relationships. In this case, the representations that need to be compared could be generated in cerebral cortex and relayed to the cerebellum through the inferior olive. In the second one, two parallel closed loops facilitate this process by providing different sources of information to different subregions of the cerebellum and the inferior olive. The inferior olive could then integrate information across these two loops to provide the error input to the cerebellum, which then uses this information to retune the mapping between two representations in cerebral cortex. In the third pathway, two converging closed loops facilitate this process, by providing two different sources of information to one cerebellar region. This region could then project back to cerebral cortex to alter one of the representations.

Our results favor a model in which adaptation is driven by lexico-semantic access that forms a bridge into sensorimotor representations (based on our functional connectivity results that show Significant Crus I Spt functional correlations). Future studies could examine different behavioral manipulations, such as the impact of different types of distortions and degrees of distortion severities, as well as different training paradigms (e.g different types of feedback). In addition, technical advances (such as high resolution functional magnetic resonance imaging and high definition track tracing methods) could be used as ways to explore the topographic organization and anatomical connectivity of the cerebellum in greater detail.

6.0 BIBLIOGRAPHY

Ackermann, H., Graber, S., Hertrich, I., & Daum, I. (1997). Categorical speech perception in cerebellar disorders. *Brain and Language , 60*, 323-331.

Ackermann, H., Mathiak, K., & Ivry, R. B. (2004). Temporal organization of "internal speech" as a basis for cerebellar modulation of cognitive functions. *Behavioral and cognitive neuroscience reviews , 3* (1), 14-22.

Adank, P., & Devlin, J. T. (2009). On-line plasticity in spoken sentence comprehension: adapting to time-compressed speech. *Neuroimage* .

Albus, J. (Math Biosci). A theory of cerebellar function. *10*, 25-61.

Allen, G., McColl, R., Barnard, H., Ringe, W. K., Fleckenstein, J., & Cullum, M. C. (2005). Magnetic resonance imaging of cerebellar-prefrontal and cerebellar-parietal functional connectivity. *NeuroImage , 28*, 39-48.

Altmann, G. T., & Young, D. (1993). Factors affecting adaptation to time-compressed speech.

Baddeley, Lewis, & Vallar. Working memory: Looking back and looking forward. *Nature Reviews Neuroscience , 4*, 829-839.

Baizer, J., & Glickstein, M. (1974). Proceedings: Role of cerebellum in prism adaptation. *J Physiol , 236* (1), 34P-35P.

Baizer, J., Kralj-Hans, I., & Glickstein, M. (1999). Cerebellar lesions and prism adaptation in macaque monkeys. *J Neurophysiol* , 1960-5.

Bloedel, J. R., & Bracha, V. (1998). Current concepts of climbing fiber function. *The Anatomical Record* .

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *NeuroImage , 32*, 821-841.

Booth, J. R., Wood, L., Lu, D., Houk, J., & Bitan, T. (2006). The role of the basal ganglia and cerebellum in language processing. *Brain Research* .

Braitenberg, V. (1967). Is the cerebellar cortex a biological clock in the millisecond range? *Prog Brain Res , 25*, 334-346.

Caesar, K., Thomsen, K., & Lauritzen, M. (2003). Dissociation of spikes, synaptic activity, and activity-dependent increments in rat cerebellar blood flow by tonic synatptic inhibition. *PNAS , 100* (26), 16000-16005.

Callan, A. M., Callan, D. E., Tajima, K., & Akahane-Yamada, R. (2006). Neural processes involved with perception of non-native durational contrasts. *Neuroreport , 17* (12), 1353-1357.

Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage , 22*, 1182-1194.

Callan, D. E., Tajima, K., Callan, A. M., KuboRieko, Masaki, S., & Akahane-Yamada, R. (2003). Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *NeuroImage , 19*, 113-124.

Carreiras, M., Seghier, M., Baquero, S., Estevez, A., Lozano, A., Devlin, J., et al. (2009). An anatomical signature for literacy. *Nature , 461* (7266), 983-986.

Catz, N., Dicke, P., & Thier, P. (2008). Cerebellar-dependent motor learning is based on pruning a Purkinje cell population response. *PNAS , 105* (20), 7309-7314.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Ann. N.Y. Acad. Sci. , 1169*, 15-34.

Clower, D. M., Hoffman, J. M., Votaw, J. R., Faber, T. L., Woods, R. P., & Alexander, G. E. (1996). Role of posterior parietal cortex int he recalibration of visually guided reaching. *Letters to Nature .*

Daum, I., & Ackermann, H. (1995). Cerebellar contributions to cognition. *Behavioural Brain Research , 67*, 201-210.

Davis, M. H., & Johnsrude, I. (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience , 23* (8), 3423-3431.

Davis, M. H., Hervais-Adelman, A., Taylor, K., Johnsrude, I. S., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental psychology: General , 134* (2), 222-241.

Desai, R., Liebenthal, E., Waldron, E., & Binder, J. R. (2008). Left posterior temporal regions are sensitive to auditory categorization. *Journal of Cognitive Neuroscience .*

Desmond, J. E., & Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: language, learning and memory. *Trends in Cognitive Sciences , 2* (9), 335-362.

Desmond, J. E., Gabrieli, J. D., Wagner, A. D., Ginier, B. L., & Glover, G. H. (1997). Lobular patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by functional MRI. *The Journal of Neuroscience , 17* (24), 9675-9685.

Desmond, J., Chen, S., & Shieh, P. (2005). Cerebellar transcranial magnetic stimulation impairs verbal working memory. *Ann Neurol , 58* (4), 554-560.

Diedrichsen, J., Hashambhoy, Y., Rane, T., & Shadmehr, R. (2005). Neural correlates of reach errors. *The Journal of Neuroscience , 25* (43), 9919-9931.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Cuurent Opinion in Neurobiology* , 732-739.

Doyon, J., Song, A. W., Karni, A., Lalonde, F., Adams, M. M., & Ungerleider, L. G. (2002). Experience-dependent changes in cerebellar contributions to motor sequence learning. *PNAS* , *99* (2), 1017-1022.

Dum, R. P., & Strick, P. L. (2003). An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *Journal of Neurophysiology* , 634-639.

Egan, J. P. (1948). Articulation Testing Methods. *The Laryngoscope , 58,* 955-991.

Ethier, V., Zee, D., & Shadmehr, R. (2008). Changes in control of saccades during gain adaptation. *J Neurosci , 28* (51), 13929-13937.

Fenn, K. M. (2003). Consolidation during sleep of perceptual learning of spoken language. *Letters to Nature* .

Fesl, G., Bruhns, P., Rau, S., Wiesmann, M., Ilmberger, J., Kegel, G., et al. (2009). Sensitivity and reliability of language laterality assessment with a free reversed association task-a fMRI study. *Eur Radiol* .

Fiez, J. A. (1996). Cerebellar contributions to cognition. *Neuron , 16*, 13-15.

Fiez, J. A., Petersen, S. E., Cheney, M. K., & Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage. *Brain , 115*, 155-178.

Fiez, J. A., Raichle, M. E., Balota, D. A., Tallel, P., & Petersen, S. E. (1988). PET activation of posterior temporal regions during auditory word presentation and verb generation. *Nature , 331*.

Fox, M., & Raichle, M. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neursoci , 8* (9), 700-711.

Francis, A., Nusbaum, H., & Fenn, K. (2007). Effects of training on the acoustic phonetic represnetation of synthetic speech. *J Speech Lang Hear Res , 50* (6), 1445-1465.

Gaab, N., Gabrieli, J. D., & Glover, G. H. (2007). Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. *Human Brain Mapping , 28,* 703-720.

Galantucci, B., Fowler, C. A., & Turvey, M. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review , 13* (3), 361-377.

Ganong, W. I. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance , 6* (1), 110-125.

Gebhart, A. L., Petersen, S. E., & Thach, W. (2002). Role of the posterolateral cerebellum in language. *Annals of the New York Academy of Sciences , 978,* 318-333.

Ghez, C., & Thach, T. W. (2000). *Principles of Neural Science* (Fourth Edition ed.). (E. R. Kandel, J. H. Schwartz, & T. M. Jessell, Eds.) New York: McGraw-Hill Companies, Inc.

Glickstein, M. (2007). What does the cerebellum really do? *Current Biology* , R824-7.

Greenspan, S. L., Nusbaum, H. C., & Pisoni, D. B. (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition , 14* (3), 421-433.

Grodd, W., Hulsmann, E., Lotze, M., Wildgruber, D., & Erb, M. (2001). Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. *Human Brain Mapping , 13,* 55-73.

Guenther, F. H., & Ghosh, S. S. (2003). A model of cortical and cerebellar function in speech. *Proceedings of the XVth International Congress of Phonetic Sciences.* Barcelona: 15th ICPhS Organizing Committee.

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

Haarmeier, T., & Thier, P. (2007). The attentive cerebellum - myth or reality? *The cerebellum , 6,* 177-183.

Habas, C. (2009). Functinoal imaging of the deep cerebellar nuclei: a review. *Cerebellum* .

Heim, S., Eickhoff, S. B., Friederici, A. D., & Amunts, K. (2009). Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. *Brain Struct Funct , 213,* 441-456.

Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. Perceptual learning of noise vocoded words: effects of feedback and lexicality.

Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance , 34* (2), 460-474.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition , 92*, 67-99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Neuroscience Reviews , 8*, 393-402.

Hoshi, E., Tremblay, L., Feger, J., Carras, P. L., & Strick, P. L. (2005). The cerebellum communicates with the basal ganglia. *Nature Neuroscience* .

Houde, J. F., & Jordan, M. I. (1988). Sensorimotor adaptation in speech production. *Science , 279* (5354), 1213-1216.

Imamizu, H., Kuroda, T., Miyauchi, S., Yoshioka, T., & Kawato, M. (2003). Modular organization of internal models of tools in the human cerebellum. *Proc Natl Acad Sci USA , 100* (9), 5461-6566.

Indefrey, P., & Levelt, W. (2004). The spatial and temporal signatures of word production components. *Cognition , 91* (1-2), 101-144.

Ito, M. (2006). Cerebellar circuitry as a neuronal machine. *Progress in Neurobiology* , 272-303.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Neuroscience Reviews* , 1-10.

Ivry, R. B. (2004). The neural representation of time. *Current Opinion in Neurobiology , 14*, 225-232.

Ivry, R. (1996). The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology* , 851-857.

Ivry, R., Keele, S., & Diener, H. (1988). Dissociation of the lateral and medial cerebellum in movement timing and movement execution. *Exp Brain Res* , 167-180.

Jenmalm, P., Schmitz, C., Forssberg, H., & Ehrsson, H. (2006). Lighter or heavier than predicted: neural correlates of corrective mechanisms during erroneously programmed lifts. *The Journal of Neuroscience , 26* (35), 9015-9021.

Joanisse, M. F., Zevin, J. D., & McClandliss, B. D. (2007). Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cerebral Cortex , 17,* 2084-2093.

Jones, J. A., & Munhall, K. (2003). Learning to produce speech with an altered vocal tract: the role of auditory feedback. *J. Acoust. Soc Am. , 113* (1), 532-543.

Jones, J. A., & Munhall, K. (2000). Perceptual calibration of F0 production: evidence from feedback perturbation. *J Acoust Soc Am. , 108* ((3 Pt 1)), 1246-1251.

Jones, J. A., & Munhall, K. (2005). Remapping auditory-motor representations in voice production. *Current Biology , 15,* 1768-1772.

Kawato, M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biol Cybern* , 95-103.

Kawato, M., & Gomi, H. (1993). The cerebellum and VOR/OKR learning models. *Trends Neurosci , 16* (5), 177-178.

Kelly, R. M., & Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *The journal of neuroscience , 23* (23), 8432-8444.

Krienen, F. M., & Buckner, R. L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cerebral Cortex* .

Laforce, R. J., & Doyon, J. (2002). Differential role for the striatum and cerebellum in response to novel movements using a motor learning paradigm. *Neuropsychologia , 40,* 512-517.

Leech, R., Holt, L. L., Devlin, J. T., & Dick, F. (2009). Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *The Journal of Neuroscience , 29* (16), 5234-5239.

Leiner, H., Leiner, A., & Dow, R. (1987). Cerebro-cerebellar learning loops in apes and humans. *Ital. J. Neurol. Scil , 8,* 425-436.

Li, T., & Fu, Q.-J. (2006). Perceptual adaptation to spectrally shifted vowels: training with nonlexical labels. *JARO* .

Liss, J. M., Spitzer, S. M., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *J. Acoust. Soc. Am. , 112* (6), 3022-3030.

Logothetis, N. K., & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magnetic Resonance Imaging , 22,* 1517-1531.

Logothetis, N., & Wandell, B. (2004). Interpreting the BOLD signal. *Annu Rev Physiol* , 735-769.

Logothetis, N., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature , 412*, 150-157.

Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences , 13* (3), 110-114.

Magnuson, J., & Nusbaum, H. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *J Exp Psychol Hum Percept Perform , 33* (2), 391-409.

Marr, D. (1969). A theory of cerebellar cortex. *J Physiol , 202* (2), 437-470.

Mathiak, K., Hertrich, I., Grodd, W., & Ackermann, H. (2002). Cerebellum and speech perception: a functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience , 14* (6), 902-912.

Mathiak, K., Hertrich, I., Grodd, W., & Ackermann, H. (2004). Discrimination of temporal information at the cerebellum: functional magnetic resonance imaging of nonverbal auditory memory. *NeuroImage , 21* (1), 154-162.

Mauk, M., Medina, J., Nores, W., & Ohyama, T. (2000). Cerebellar function: Coordination, learning or timing? *Current Biology , 10*, R522-R525.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: lexical adaptation to a novel accent. *Cognitive Science* , 543-562.

McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception. *TRENDS in Cognitive Sciences , 8* (10), 363-369.

McCormick, D., & Thompson, R. (1984). Cerebellum: essential involvement in the classically conditioned eyelid response. *Science , 223* (4633), 296-299.

McGurk, H., & McDonald, J. (1976). Hearing lips and seeing voices. *Nature , 264*, 746-747.

McQueen, J. M., Norris, D., & Cutler, A. (2006). Are there really interactive processes in speech perception? *TRENDS in Cognitive Sciences , 10* (12), 533.

Miall, R., & Jenkinson, E. (2005). Functional imaging of changes in cerebellar activity related to learning during a novel eye-hand tracking task. *Exp Brain Res , 166*, 170-183.

Middleton, F. A., & Strick, P. L. (1998). Cerebellar output: motor and cognitive channels. *Trends in Cognitive Sciences , 2* (9), 348-354.

Myers, E. B., & Blumstein, S. E. (2007). The neural bases of the lexical effect: an fMRI investigation. *Cerebral Cortex* .

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual Learning in Speech. *Cognitive Psychology , 47,* 204-238.

Obleser, J., Wise, R. J., Dresner, A. M., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience , 27* (9), 2283-2289.

Ohyama, T., Nores, W. L., Murphy, M., & Mauk, M. D. (2003). What the cerebellum computes. *TRENDS in Neuroscience* , 222-227.

Okada, K., & Hickok, G. (2006). Identification of lexical-phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. *NeuroReport* , 1293-1296.

O'Reilly, J. x., Beckman, C. F., Tomassini, V., Ramnani, N., & Johansen-Berg, H. (2009). Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. *Cerebral Cortex* .

Papathanassiou, D., Mellet, E. E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. A common language network for comprehension and production: a contribution to the definition of the language epicenters with PET. *NeuroImage , 11*, 347-357.

Penhune, V., & Doyon, J. (2005). Cerebellum and M1 interaction during early learning of timed motor sequences. *Neuroimage , 26* (3), 801-812.

Petacchi, A., Laird, A., Fox, P., & Bower, J. (2005). Cerebellum and auditory function: an ALE meta-analyss of functional neuroimaging studies. *Human Brain Mapping , 25* (1), 118-128.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience , 1*, 153-170.

Pisella, L., Michel, C., Grea, H., Tilikete, C., Vighetto, A., & Rossetti, Y. (2004). Preserved prism adaptation in bilateral optic ataxia: strategic versus adaptive reaction to prisms. *Exp Brain Res* , 399-408.

Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenic, M., et al. (n.d.). Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI.

Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenic, M., et al. (2008). Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI.

Prabhakaran, R., Blumstein, S. E., Myers, E. B., Hutchison, E., & Britton, B. (2006). An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia , 44*, 2209-2221.

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *PNAS , 103* (20), 7865-7870.

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel format frequency: evidence from real-time formant manipulation. *J. Acoust. Soc. Am. , 120* (2), 966-977.

Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nature Neuroscience Reviews* , 511-522.

Research Imaging Center UTHSCSA. (2003-2009). *BrainMap Sleuth 1.2*. Retrieved from http://www.brainmap.org

Riecker, A., Kassubek, J., Groschel, K., Grodd, W., & Ackermann, H. (2006). The cerebral control of speech tempo: opposite relatinoship between speaking rate and BOLD signal changes at striatal and cerebellar structures. *NeuroImage , 29*, 46-53.

Robertson, E., & Miall, R. (1999). Visuomotor adaptation during inactivation of the dentate nucleus. *NeuroReport , 10* (5), 1029-1034.

Schmahmann, J. D., Doyon, J., McDonald, D., Holmes, C., Lavoie, K., Hurwitz, A. S., et al. (1999). Three-dimensional MRI atlas of the human cerebellum in proportional stereotaxic space. *NeuroImage* , 233-260.

Schmahmann, J., & Pandya, D. (1991). Projections to the basis pontis from the superior temporal sulcus and superior temporal region in the rhesus monkey. *Journal of Comparative Neurology* , 224-248.

Schmidt, C. F., Zaehle, T., Meyer, M., Geiser, E., Boesiger, P., & Jancke, L. (2008). Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. *Human Brain Mapping , 29*, 46-56.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-prime user's guide. Pittsburgh: Psychological Software Tools Inc.

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* , 473-500.

Scott, S., Rosen, S., Lang, H., & Wise, R. (2006). Neural correlates of intelligibility in speech investigated with noise vocoded speech-a positron emission tomography study. *J Acoust Soc Am* , 1075-83.

Shannon, R., Zeng, F., Kamath, V., Wygonski, J., & Ekelid, M. Speech recognition with primarily temporal cues. *Science , 270*, 303-304.

Shmuel, A., Augath, M., Oeltermann, A., & Logothetis, N. (2006). Negative functional MRI response correlates with decreaseds in neuronal activity in monkey visual area V1. *Nature Neuroscience , 9* (4), 569-577.

Soetedjo, R., Kojima, Y., & Fuchs, A. (2008). Complex spike activity signals the direction and size of dysmetric saccade errors. *Prog Brain Res , 171*, 153-159.

Soros, P., Sokoloff, L. G., Bose, A., McIntosh, A. R., Graham, S. J., & Stuss, D. T. (2006). Clustered functional MRI of overt speech production. *NeuroImage , 32*, 376-387.

Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *NeuroImage , 44*, 489-501.

Stoodley, C. J., & Schmahmann, J. D. (2009). The cerebellum and language: evidence from patients with cerebellar degeneration. *Brain and Language* .

Strick, P. L., Dum, R. P., & Fiez, J. A. (2009). Cerebellum and nonmotor function. *Annu. Rev. Neurosci.* , 413-4343.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: an approach to medical cerebral imaging.* Thieme, Germany: Stuttgart.

Thach, W. (2007). On the mechanism of cerebellar contributions to cognition. *The Cerebellum , 6*, 163-167.

Thach, W. T. (1998). What is the role of the cerebellum in motor learning and cognition? *Trends in Cognitive Sciences , 9* (2), 331-337.

Thompson, R., & Steinmetz, J. (2009). The role of the cerebellum in classical conditioning of discrete behavioral responses. *Neuroscience , 162*, 1732-755.

Thomsen, K., Offenhauser, N., & Lauritzen, M. (2004). Principal neuron spiking: neither necessary nor sufficient for cerebral blood flow in rat cerebellum. *J Physiol* , 181-189.

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

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

Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., et al. (2007). Neural correlates of auditory feedback control in human. *Neuroscience , 146*, 499-503.

Vaillancourt, D. E., Thulborn, K. R., & Corcos, D. M. (2003). Neural basis for the processes that underli visually guided and internally guided force control in humans. *Neurophysiol , 90*, 3330-3340.

van de Ven, V., Esposito, F., & Christoffels, I. K. (2009). Neural network of speech monitoring overlaps with overt speech production and comprehension networks: A sequential spatial and temporal ICA study. *NeuroImage , 47*, 1982-1991.

Voogd, J., & Glickstein, M. (1998). The anatomy of the cerebellum. *Trends Neurosci , 9*, 370-375.

Ward, D. B. (2000). *Simultaneous Interference for FMRI Data.* Milwaukee, WI: Biophysics Research Institute Medical College of Wisconsing.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience , 7* (7), 701-702.

Wolpert, D. M., Miall, C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Congitive Sciences , 2* (9), 338-347.

Woods, R., Cherry, S., & Mazziotta, J. (1992). Rapid automated algorithm for aligning and reslicing PET images. *J Comput Assist Tomogr , 16* (4), 620-633.

Woods, R., Mazziotta, J., & Cherry, S. (1993). Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography , 17*, 536-546.

Xian, H.-D., Fontejin, H. M., Norris, D. G., & Hagoort, P. (2009). Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral Cortex* .

Yamamoto, K., Hoffman, D., & Strick, P. (2006). Rapid and long-lasting plasticity of input-output mapping. *J Neurophysiol , 96* (5), 2797-2801.

Yamazaki, T., & Tanaka, S. (2009). Computational models of timing mechanisms in the cerebellar granular layer. *Cerebellum* .

Zaehle, T., Schmidt, C. F., Meyer, M., Baumann, S., Baltes, C., Boesiger, P., et al. (2007). Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. *NeuroImage , 37*, 1195-1204.

Zatorre, R., Evans, A., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science , 256*, 846-849.

Zeng, F. (2004). Trends in cochlear implants. *Trends Amplif , 8* (1), 1-34.