Audiovisual Integration in Apraxia of Speech: EEG Evidence for Processing Differences

By

Melissa Randazzo

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

Columbia University

2016

Abstract

Audiovisual Integration in Apraxia of Speech: EEG Evidence for Processing Differences

Melissa Randazzo

Speech perception is a unique audiovisual experience in part because timing of the speech signal is influenced by simultaneous overlapping gestures in coarticulation. Apraxia of speech (AOS) is a motor planning disorder that impairs coarticulation. Imaging studies show that brain regions damaged in AOS are critical to audiovisual speech perception. Although AOS is a motor planning disorder, individuals with AOS may have a disruption to the perceptual system for speech gestures. To evaluate this hypothesis we investigated audiovisual mismatch negativity (MMN) brain responses in adults with damage to Broca's area (n =5) compared to a healthy age-matched comparison group (n = 5). We utilized the McGurk effect, in which incongruent auditory and visual information alters perception. Participants viewed videos of a speaker articulating the syllable /ba/ (standard) for 80% trials and /ga/ (deviant) for 20% of the trials while the auditory stimulus /ba/ remained consistent throughout. Responses to this McGurk audiovisual condition were compared to an inverse McGurk audiovisual condition in which the visual stimulus remained constant while the auditory stimulus changed, and a visual-only condition without sound to control for evoked activity from changes to the visual stimulus.

Incongruent McGurk deviants elicited an MMN over left hemisphere electrodes in the comparison group, while the AOS group exhibited a later, attention-based response, a P300. The AOS group similarly responded to inverse McGurk deviants, which do not require fusion of the percept, with a P300 response, indicating that auditory and visual aspects of the incongruent McGurk deviants were not integrated. In the visual-only control condition, the AOS group showed a left-lateralized MMN, suggesting greater influence of visual processing when confronted with conflicting multisensory information compared to the comparison group. Overall, the comparison group's responses were indicative of early and automatic audiovisual integration of incongruent McGurk percepts while the responses of the AOS group showed contributions of both attentional and visual processing. The timing of the response in the AOS group was correlated with speech production characteristics of apraxia, as well as performance on taxing motor speech tasks. Results of this study support the hypothesis that AOS is a disorder beyond motor planning, with implications for higher-level linguistic and cognitive systems.

Table of Contents

List of Figures

List of Tables

Acknowledgments

This dissertation and completion of my doctoral studies would not have been possible without the support and contributions of my incredible family, friends, and colleagues.

I am forever grateful to my friends and colleagues in the Neurocognition of Language Lab (Trey Avery, Felicidad Garcia, Heather Green, Dayna Moya, Lisa Levinson, Chaille Maddox, Guannan Shen, Lauren Shuffrey, Laura Sanchez, and Paula Garcia) for their help, guidance, and support these past few years. Your comradery, generosity, and dedication makes me so proud to have been part of this lab.

This work would not have been possible without the support of our dedicated staff in the Department of Biobehavioral Sciences- Maria Lamadrid, Erynn Lowery, and Yvonne Wallace. Thank you for all you have done to support me over the years.

Special thanks to my committee members- Dr. Lisa Edmonds, Dr. Michelle Troche, Dr. Robert Remez, Dr. Reem Khamis-Dakwar, and Dr. Karen Froud.

The many participants, aphasia survivors, and co-survivors, who dedicated their time and energy to help with this project- I thank you from the bottom of my heart.

Dr. Reem Khamis-Dakwar, thank you for your mentorship, support, and cheerleading over the past few years. I am looking forward to building great things with you in the years to come.

My husband, Jason Wagner, thank you for sticking with me in the toughest of times along this journey. My parents, Marie and John Randazzo, for your patience and support during my long career as a student.

Most of all, Dr. Karen Froud, this work would not have been possible without you. You have been the greatest mentor and inspiration for the academic, mother, and human being I want to be. Thank you for all you have done to make this dream possible for me and for all you have taught me over the years.

Dedication

This work is dedicated-

To Harvey Alter and Michael Young, who dedicated themselves to serving the aphasia community in NYC. Your community taught me to be the best clinician I can be. The impact of what you have built will live on. You are sorely missed and are forever in our hearts.

To the work and career of A. Damien Martin, whom I never had the chance to meet, but much admire.

To the survivors and co-survivors of aphasia, you are my inspiration to continue this work.

To my son, Mason Edward Wagner, you are my reason to push myself further every day.

Chapter I

INTRODUCTION

Speech perception is an inherently multisensory, or audiovisual process, in which viewing the speaker's articulatory movement influences or enhances perception (McGurk & MacDonald, 1976; Sumby & Pollack, 1954). Evidence for the audiovisual (AV) nature of speech perception comes from the McGurk Effect, which shows that incongruent auditory and visual information alters perception (McGurk & MacDonald, 1976). Previous EEG studies of the McGurk effect have shown a mismatch negativity (MMN) event related potential (ERP) response to incongruent audio-visual representations (Colin, Radeau, Soquet, Demolin, Colin, & Deltenre, 2002; Colin, Radeau, Soquet, & Deltenre, 2004; Hessler, Jonkers, Stowe, & Bastiaanse, 2013; Musacchia, Sams, Nicol, & Kraus, 2006; Sams et al, 1991).

Apraxia of speech (AOS) is an aging-related disorder that impedes verbal communication. Several fMRI studies have shown that a speech-production motor network, including sites damaged in AOS, is critical to audiovisual speech perception (Ojanen et al., 2005; Skipper, Nusbaum, & Small, 2005). To investigate whether there is a breakdown in audiovisual integration in AOS, this dissertation study exploits the McGurk effect using electroencephalography (EEG), a measure of neuronal communication that permits the derivation of Event-Related Potentials (ERPs). ERPs can provide "signatures" of various aspects of cognitive processing, including processing and recognition of speech sounds (Luck, 2005). The MMN (MisMatch Negativity) is one such ERP, constituting an early and automatic response to speech sounds, which makes it an ideal outcome measure for investigating online perceptual processing in individuals whose behavioral responses would be limited by linguistic and motoric impairments (Näätänen, 1982; Näätänen, 1990; Näätänen et al., 2012). The goal of this dissertation is to examine neurophysiological responses of individuals with AOS in order to determine if they show audiovisual processing differences compared to healthy adults, and further, if these neurophysiological responses are related to speech production features of AOS. Knowledge regarding disruption to audiovisual processing in AOS addresses critical barriers in understanding the nature of motor speech impairments, accurate diagnostic criteria, and effective treatments.

This dissertation is organized as follows: the rest of Chapter I provides a general introduction to the background and significance, as well as the innovation, of exploring audiovisual integration in AOS. Chapter II provides a description of AOS, including the evolving conceptualization of the disorder in relation to different theoretical perspectives. Chapter III discusses audiovisual integration and introduces the McGurk effect through event-related potential studies. Chapter IV outlines the research questions and hypotheses. Chapter V describes the design of the study and the methods. Chapter VI describes the data recording and collection procedures. Chapter VII details the results of the study and provides an analysis. In Chapter VIII the implications of the results are discussed in relation to the research questions, within the context of the current literature.

1.1 Background and Significance

It is estimated that 1,000,000 Americans are living with aphasia, the debilitating language disorder caused by stroke, and an additional 80,000 acquire the disorder each year (National Aphasia Association, 2016). Aphasia has been shown to have a more negative impact on quality of life than cancer or Alzheimer's Disease (Lam & Wodchis, 2010). Apraxia of speech (AOS) is a sensorimotor speech disorder that frequently co-occurs with aphasia. AOS also occurs in the neurodegenerative disorder primary progressive aphasia (PPA) and its presence or absence is the main characteristic utilized for differential diagnosis between sub-types of PPA (Croot, Ballard, Leyton, & Hodges, 2012). The incidence of stroke and therefore AOS is expected to rise with the continued aging of the population, with an estimated projected annual healthcare cost of $34 billion (Mozaffarian et al., 2015).

Cases of pure AOS in the absence of aphasia occur from damage to left premotor cortex and precentral gyrus, plus Brodmann Area 44 (BA 44) of Broca's area in cases with concomitant aphasia (Graff-Radford et al., 2014; Hillis et al., 2004). Descriptions of AOS include phonetic impairments in the planning, sequencing, coordination, and timing of articulatory movements (Code, 2005; Darley & Aronson, 1975; Kent & Rosenbek, 1983). Patients with AOS present with slow, effortful speech that is produced with frequent sound errors and disturbed prosody (Strand, Duffy, Clark, & Josephs, 2014; Wambaugh et al., 2006). Such errors are characterized as difficulty with coarticulation, or the articulation of conceptually distinct speech sounds together such that one influences the other (Hardcastle & Hewett, 2006). Although AOS most frequently co-occurs with a linguistic aphasic impairment, it is largely considered a deficit in speech motor planning. Several behavioral studies have shown semantic and phonological involvement in unisensory perceptual responses in AOS (Maas, Barlow, Robin, & Shapiro, 2002; Maas, Gutierrez, & Ballard, 2014; Strand & McNeil, 1996;). Evidence for linguistic involvement in AOS indicates the potential for involvement of multiple sensory systems and higher-level representational mechanisms.

The designation of AOS as a disorder of motor planning is historically based on behavioral observations of speech production deficits. While speech production is clearly impacted in AOS, the possibility of an underlying perceptual deficit has been relatively ignored in the research literature. Current models of motor speech production and perception acknowledge that speech is a multisensory process, comprised of auditory, visual, motor, and

proprioceptive components (e.g., Hickok 2012, 2013, 2014; Pickering & Garrod, 2013; Tian & Poeppel, 2012). Converging evidence from functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) supports the notion that speech motor areas are implicated in speech perception. Specifically, fMRI work shows that the posterior portion of Broca's area is activated while viewing speech motor movements (Ojanen et al., 2005; Skipper, Nusbaum, & Small, 2005). Moreover, work on somatosensory processing has shown that alterations or limitations to jaw and lip movements adjust phoneme boundaries during categorical perception, and conversely, that alterations to auditory feedback can impact production of fundamental frequency (Lametti, Nasir, & Ostry, 2012; Nasir & Ostry, 2006). These results suggest that the motor system is recruited in mapping acoustic input into phonetic code via articulatory gestures. Thus it is plausible that individuals with speech motor impairments may have a breakdown in perceiving this information.

While research across disciplines has shown that speech production and perception are coupled processes, clinical speech language pathology continues to draw a sharp distinction between individuals who have speech production disorders and those with representational phonological deficits. In the case of AOS, it is presumed that there is a breakdown in the motor planning or programming phase of speech production, yielding phonetic errors. This is in contrast to assumed phonological, or abstract representational, errors that occur in aphasia and language-based disorders. However, current theories of speech and language do not identify these issues as mutually exclusive, but rather inter-related (Hickok & Poeppel, 2004; Hickok, 2012; Libermann & Mattingly, 1985; Tourville & Guenther, 2011). This false dichotomy in the clinical domain has polarized the conceptualization of motor speech disorders and linguistic impairments. The nature of AOS has been debated in the literature, and some of the outstanding questions regarding this disorder are closely tied to its resistance to improvement following treatment. This dissertation study has the potential to provide fundamental information about speech perception in motor planning disorders, with implications for theories of motor planning, diagnostic criteria, and treatment of AOS.

1.2 Innovation

The current study provides theoretical and methodological innovations to approaches in understanding the nature of the impairment in AOS. Research in this area has been limited to analysis of speech production or measurement of speech perception via reaction times. For example, one study to date demonstrated that patients with AOS do not benefit from bimodal AV cueing in speech perception tasks, and this difficulty is specifically related to speech but not nonspeech stimuli (Ziegler & Schmid, 2006). Using such methods, however, the presence of a perceptual deficit cannot be disambiguated from a production deficit (Libermann & Mattingly, 1985), and therefore speech production analysis is inadequate to inform our understanding of AOS. Investigations of speech perception in AOS have relied on reaction times, which is problematic for patients with AOS, who often suffer from general motoric impairments that impede responses. Moreover, investigations of speech perception in AOS to date have focused on unisensory responses to auditory stimuli (e.g., Maas, Barlow, Robin, & Shapiro, 2002; Maas, Gutierrez, & Ballard, 2014; Strand & McNeil, 1996), ignoring the audiovisual nature of speech perception.

Patients with AOS typically have damage to Broca's area, which is integral to speech production. Neuroimaging studies have shown that Broca's area is involved in audiovisual speech perception in healthy adults. Specifically, these studies have shown that Broca's area is integral to cognitive conflict resolution when confronted with incongruent audiovisual speech stimuli (Ojanen et al., 2005). These findings have been debated, as there is evidence that Broca's area shows greater responses to visual-only stimuli than to audiovisual stimuli (Matchin et al., 2014). Demonstration of an audiovisual impairment in patients with AOS who have suffered damage to Broca's area would clarify whether Broca's area is involved in audiovisual speech perception. This information would provide additional insight into characteristics of AOS and information for differential diagnosis of PPA sub-classification.

Demonstration of an audiovisual impairment for speech perception in AOS would also support theories of linguistic involvement in an assumed motor planning disorder. There is some evidence for linguistic involvement in AOS, such as semantic facilitation of motor production (Buxbaum & Saffran, 2001), and findings indicating that training phonologically complex targets yields greater treatment gains (e.g., Maas et al., 2002). Some evidence shows that AOS is associated with disturbances in linguistic representation, such as transfer between phonological

and phonetic representations (Galluzi, Baureca, Guariglia, & Romani, 2015) or difficulties with phonological encoding (Laganaro, 2012). Recent fMRI work has shown that brain regions damaged in AOS overlap with those implicated in verbal short-term memory (Hickock, Rogalsky, Chen, et al., 2014). Evidence for multisensory processing deficits in adults with AOS would support conceptualizations of AOS as a disorder beyond motor planning, including higher-level cognitive and linguistic skills.

This dissertation utilizes electroencephalography (EEG), which allows for online measurement of multisensory perception in the absence of overt physical response. AV integration in AOS has been researched via behavioral and neuroimaging methods. Investigation utilizing EEG further contributes to this literature by specifically examining the time course of brain activations related to multisensory processing in AOS. The next chapter provides an overview of AOS as a diagnostic entity and its conceptualization within current models of speech processing and production.

Chapter II

APRAXIA OF SPEECH

2.1 Clinical Description and Diagnostic Markers

AOS is an acquired neurogenic speech disorder that impacts the ability to plan sensorimotor commands for speech. AOS can occur in the absence of, or in addition to, muscular weakness associated with dysarthria and/or linguistic impairments associated with aphasia (Duffy, 2013). The main etiology of AOS is cerebrovascular accident, although it may be associated with progressive neurological diseases in subtypes of primary progressive aphasia or corticobasal degeneration (Duffy & Josephs, 2012; Rosenfeld, 1991). AOS caused by cerebrovascular accident, the focus of the current study, is typically caused by carotid system thromboses that generally lodge in the left middle cerebral artery (MCA), resulting in non-hemorrhagic stroke.

The term "apraxia" was applied to this speech disorder in 1969 by Darley and colleagues at the Mayo Clinic, and later specified as an impairment in planning the sequential positions of the speech musculature (Darley et al., 1975). Subsequent definitions added descriptions of deficiencies in both sequential and temporal aspects of speech motor control (Kent & Rosenbeck, 1983). The temporal and sequential planning impairments are manifested in both phonetic and prosodic speech disturbances. The core features of AOS include: effortful speech with trial and error groping during self-correction attempts; prosodic abnormalities that affect rate, rhythm, and stress; inconsistent productions of the same utterance upon repetition; and difficulty initiating utterances (Ogar, Slama, Drokners, Amici, Luisa Gorno-Tampini, 2005). Phonetic errors in AOS typically involve place of articulation, most often with fricatives, affricates, and consonant clusters (Duffy, 2013). Prosodic abnormalities are thought to be secondary to articulation errors, as rate of speech is slowed in anticipation of errors (Darley & Aronson, 1975). Speech rate is also slowed by equal stress placement in sentences, lengthened vowels, and pauses between syllables and words (Duffy, 2013; Seddoh et al., 1996; Strand & McNeil, 1996).

2.2 AOS in Models of Speech Motor Control

Historically, conceptualization of AOS has been based on theoretical models of speech motor control within the field of speech-language pathology. Models proposed by Darley, Aaronson, and Brown (1975a,b), and subsequently van der Merwe (1997), are rooted in observations of patient characteristics, with origins in the field of aphasiology. Darley et al. (1975a) proposed a three-stage model to characterize the breakdown of the speech musculature during apraxic speech production in the absence of muscle weakness or linguistic impairments. The model is comprised of the auditory speech processor (ASP), a central language processor (CLP), and the motor speech programmer (MSP). In Darley's model the three components operate in a sequential and parallel fashion, interacting with higher-level processes including conceptualization, language formulation, and motor planning and programming. Initially, the CLP encodes the cognitive and linguistic goals of the spoken message, with input from the auditory speech processor (ASP) that can modify cognitive and linguistic goals in the context of an ongoing communicative exchange, and/or self-monitoring of continuous speech. Upon establishment of the phonological representation of the speaker's message in the CLP, the MSP is activated. In this stage, upon activation the MSP translates the phonological representation into a neuromotor code specifying the parameters of movement by selecting, sequencing, activating, and controlling the preprogrammed maneuvers learned through practice. Finally, the motor plan delineated by the MSP is executed by the speech musculature, and projected to the motor cortex. Darley presumes that the breakdown in AOS occurs at the level of the MSP, where the spatiotemporal parameters of speech are determined. In this conceptualization of the impairment, the motor planning component of AOS is considered post-linguistic.

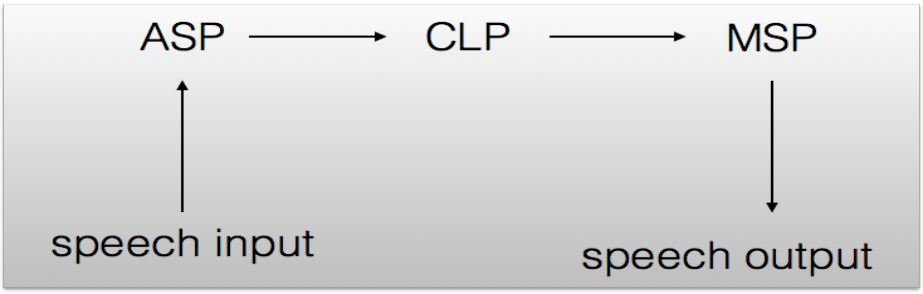


Figure 1. Illustration of Darley's (1975) Motor Speech Processor, adapted from figure in Lass (2014).

van der Merwe (1997) proposed a four-stage model of sensorimotor speech control, building upon Darley et al.'s (1975) earlier model. In the van der Merwe model first the basic linguistic units of the spoken message, phonemes, are selected. Upon selecting the required phonemes, motor planning of the overall goal is initiated to organize them into the spatiotemporal codes for speech. Subsequent to motor planning, a stage of motor programming occurs in which muscle-specific motor programs are selected and sequenced. Ultimately, the motor program sequences are executed by the speech musculature as the spoken message. van der Merwe's fourth stage of the model updated the understanding of speech disorders. A disorder of motor planning was presumed to occur due to damage in the motor cortex in the third stage, motor programming, predominantly impacting the adaptation of the articulators to the phonetic context. The distinction between motor planning and motor programming here is presented in a serial fashion, with processing outcomes moving further from linguistic involvement with each subsequent step.

Processing Level	Function
1. Linguistic Processing	specification of semantic, syntactic, and phonological aspects of speech production
2. Motor Planning	identification of motor goal that determines the motor plan
3. Motor Programming	specification of articulatory goals such as place and manner in conversion of the motor plan to program
4. Motor Execution	realization of the motor program via speech production

Table 1. Illustration of van der Merwe's (1997) four stage model, adapted from table in Ward & Scaler Scott (2011).

2.3 Neuroanatomy of AOS

Patients with AOS who have experienced a stroke typically have lesions in the anterior left hemisphere of the brain. The specific lesion site associated with AOS only is difficult to determine because in many patients AOS co-occurs with aphasia. Moreover, differences in time post-onset, diagnostic criteria, localization method, and task demands between studies have contributed to inconsistent results. Dronkers (1996) utilized a lesion overlay method to examine MRI and CT scans of patients with left hemisphere ischemic strokes in the chronic phase, with and without AOS. Findings demonstrated a dissociation between patients with AOS (n=25) and those with similar left hemisphere lesions but no speech impairment (n=19). AOS patients all had a lesion in the precentral gyrus of the left anterior insula, a region that was spared in the patients without AOS.

Hillis and colleagues (2004) argued that while the insula may be a shared lesion site among AOS patients, the lesion overlay method does not account for the reciprocal possibility that this particular lesion site did not cause the deficit. Rather, AOS may be due to hypoperfusion or an infarct to Broca's area caused by narrowing or occlusion of the left MCA, which is also related to insular damage. Hillis et al. utilized diffusion weighted imaging (DWI) and perfusion weighted imaging (PWI) to examine dysfunctional tissue in 40 patients with and 40 patients without insular damage within 24 hours of stroke onset. Results indicated that AOS was associated with structural damage or hypoperfusion in the left posterior inferior frontal gyrus (BA44) rather than the insula.

Graff-Radford and colleagues (2014) performed analyses of magnetic resonance images (MRIs) taken 1-10 days post-stroke in 7 patients: 5 with AOS and no aphasia, and 2 with equivalent aphasia. Common areas of lesion overlap between these patients included the left premotor cortex and left precentral gyrus. Recently, New and colleagues (2015) postulated that given the variability in lesion sites identified across studies, AOS may be associated with alterations to a network of previously identified regions, rather than localized to a single region. Examination of connectivity of resting state fMRI data in a network including the bilateral inferior frontal gyrus (BA44), left premotor area, and left anterior insula, demonstrated reduced connectivity between bilateral BA44, which significantly correlated with severity of AOS impairment. Additionally, results indicated negative connectivity between left premotor area and right anterior insula, particularly in patients with more severe nonverbal oral apraxia. While the

exact role of the anterior insula in speech is unclear, the authors suggest that this inverse negativity may indicate nonspeech motor compensation of the right anterior insula in patients with AOS. Thus, examination of the motor speech network revealed that regions identified in previous studies (BA44/ left inferior frontal gyrus, premotor area, anterior insula) are all involved in the pathology of AOS; however, the role of each area and their degree of involvement in pathogenesis remains unclear.

2.4 Current conceptualizations of AOS

Despite incremental advances in our understanding of the neuroanatomy and pathogenesis of AOS, its conceptualization as a diagnostic entity has been controversial among clinicians and researchers. From uncertainty regarding the diagnostic boundaries to debate about the theoretical underpinnings, research regarding AOS has been peppered with deliberation about the underlying nature and even the existence of AOS as an independent disorder.

2.4.1 Uncertainty of diagnostic boundaries

Diagnosis of AOS depends upon excluding characteristics of its close clinical neighbors, dysarthria and aphasia. While some perceptual features of the motor speech pathology in AOS and dysarthria overlap, differential diagnosis between the two syndromes generally poses minimal difficulty due to the definable etiology of each. AOS typically results from left hemisphere stroke impacting the posterior frontal lobe while most dysarthrias stem from subcortical or unilateral upper motor neuron disease. Clinically, AOS should be distinguishable from dysarthria in that volitional speech in AOS is generally more error-prone than automatic speech tasks (Duffy, 2013). Variability of speech production is considered a hallmark of AOS and generally distinguishes it from the consistently erred productions in the dysarthrias, save for hyperkinetic and ataxic dysarthrias, which tend towards variability in speech production. However, the criterion of inconsistent productions in diagnosis of AOS has been equivocated over time. Duffy (2013) maintains that consistent error types and locations are characteristic of AOS, supported by Wambaugh et al.'s (2006) *Treatment Guidelines of AOS of the Academy of Neurologic Disorders*. Straiger and colleagues (2012) noted that given the conflicting results of several studies, including their own findings of both consistent and inconsistent error types and occurrences within and between patients, the error variability is not alone sufficient for differential diagnosis of AOS.

A greater problem is posed by a diagnosis of exclusion when considering the etiological, anatomical, and perceptual overlaps between aphasia and AOS. Cases of pure AOS are rare and many patients suffer from concomitant nonfluent aphasia due to lesions in the posterior frontal lobe impacting BA44. Given that both motor speech and language disorders coexist in patients with AOS, speech-language pathologists are tasked with disambiguating phonetically based (presumed motor speech) errors from phonological (presumed linguistic) errors. If AOS is a

disorder of motor planning, associated with impaired translation of phonological representations to phonetic realizations, phonetically based errors are expected (Darley et al., 1975b; Laganaro, 2012; McNeil, Pratt, & Fossett, 2008; van der Merwe, 1997). Inherent to this difficulty is the categorical bias of human speech perception (Galluzi, Bureca, Guariglia, & Romani, 2008). The listener may miscategorize a phonetic error for a phonological error due to the nature of speech sound perception between categorical boundaries. Moreover, Laganaro (2012) cautions that realized phonetic errors may occur as a result of competing activation of phonological representations, signaling an underlying phonological impairment as the impetus for speech distortions.

2.4.2 Dichotomous theoretical frameworks

The difficult task of disambiguating a motor speech or phonetically based error from a linguistic or phonologically based error is reflective of the historical evolution of models of speech and language processing. Over forty years ago, Martin (1974), in his objection to the term *apraxia* in reference to the impairment, noted that the conceptualization was based on "outdated" and dichotomous models that separate motor realizations from phonological representations. He reflected that these ideas were born from Aten, Johns, and Darley's (1971) description of AOS as a disorder of encoding rather than decoding, while in their previous work (Johns and Darley, 1970) the linguistic disturbance in aphasia was referred to as "perceptual". More recently, Ziegler (2012) argued that a dualist tradition, born from 20th century linguistic theories (e.g., Jakobson, 1937; DeSaussure, 1967; Chomsky & Halle, 1968) that position phonological representation as abstract and thus separate from motor realization, has polarized the conception of speech and language in clinical aphasiology. Even so, Martin's (1974) essay acknowledged that the work of Jakobson and Halle (1956) and Stevens and Halle (1967) began to describe a reciprocity between the generative rules of speech production and speech perception. These ideas were echoed by Liberman's (1967) earlier work, pre-dating his influential Motor Theory of Speech Perception, positing that phonemes are perceived in reference to how they are produced by the speaker. More recent theories recognize that abstract phonological representations interact with phonetics and may be constrained by unfolding temporal aspects of speech or relationships with sensorimotor systems (Brownman & Goldstein, 1992; Gafos & Benus, 2006; Goldrick, Baker, Murphy, & Baese-Berk, 2011; Ohala, 1990; Solé, Beddor, & Ohala, 2007).

2.5 A Unifying Theoretical Framework: Hierarchical State Feedback Control Model of Speech Production

Seeking to bridge the seemingly disparate traditions of psycholinguistic and speech motor control frameworks, Hickok (2012, 2013, 2014) proposed the Hierarchical State Feedback Control Model of Speech Production (HSFC). Hickok's HSFC, and its earlier incarnation, the State Feedback Control Model of Speech Production (SFC), attempt to address the "level driven chasm" embodied by the two camps, in which psycholinguistic theories attend to higher level processes and motor control theories attend to lower level processes. HSFC is based on earlier work by Hickok and Poeppel (2000, 2004, 2007) that identified a dual processing stream for speech perception that accomplishes word learning by sound to concept mappings via a ventral stream and sound to speech gesture mappings via a dorsal stream. Within this framework, the primary area responsible for speech perception is the posterior superior temporal lobe and auditory cortical areas bilaterally, with a ventral processing stream for auditory comprehension and a dorsal stream for sub-lexical speech tasks. In development, the two streams are acquired independently and later integrated. As there is no confirmed consensus on the neural structures responsible for speech perception, this framework posits a task-specific model. While Hickok and Poeppel negate the assumption that most experimental tasks actually represent the processes involved in natural, conversational speech perception, they differentiate between tasks that require explicit attention to phoneme segments and those that require auditory comprehension. Although each type of task activates different structures, both tasks are consistently supported by the posterior superior temporal lobe. Another inference is made about an auditory-motor interface system based on the ability to repeat heard pseudowords (in the absence of semantic information). An auditory-motor interface system may exist in the inferior parietal lobe, where sound-based representations of speech in the auditory cortex interface with articulatory-based representations in the frontal cortex, providing an account of phonological working memory. Hickok's subsequent work on the HSFC model provides further detail regarding the auditory-motor interface system.

The HSFC model has not been without criticism from both linguists and motor control theorists (Hickok, 2014). In addressing critiques from both sides, Hickok urged that cross-disciplinary descriptions of seemingly distinct yet parallel phenomena in speech production

would have greater explanatory power than the traditionally distinct accounts. In the HSFC model both linguistic and motor targets are hierarchical, with auditory targets embodying higher-level sensory goals and somatosensory targets embodying lower-level phonemic goals. Auditory targets and speech motor plans are integrated via the dorsal sensorimotor processing stream. Leading with aspects of a psycholinguistic framework (see Levelt, 1983), initiation of the model begins with a conceptual representation that is fed to the lemma[1] or lexical level. The lemma level has parallel projections to both sensory and motor high-level feedback of cortical control via a loop through auditory cortex, BA 44, and the sylvian fissure at the parietotemporal boundary (Spt). This high level loop also has parallel projections to a lower-level circuit of somatosensory-cerebellum-motor cortex. Connections between the parallel projections can be excitatory or inhibitory.

The HSFC allows for both external and internal feedback. The internal forward model, as in the non-speech motor control literature, is necessary to make on-line adjustments by predictions about current and future states of motor effectors. In other models, this is generally accomplished by a post-hoc process of efference copy – the internal copying of an external, movement-producing signal that provides for feedback and modification. In HSFC, by contrast, efference copy is integrated into the model as part of the motor planning process, allowing for internal feedback. The rationale for including efference copy in the motor planning process is that it allows for inhibitory control in the event that the wrong motor plan is selected and initiated, and in instances where the correct motor plan is selected and initiated the inhibitory signal will cancel out the copied signal between prediction and detection.

Another innovation of the HSFC model is that the integral linguistic unit is the syllable rather than the phoneme. Hickok theorizes that the auditory goals are broad while the somatosensory goals fine-tune the process with context-dependent details. Phonological representations are not abstract linguistic units but rather high-level, sensory-motor representations. These syllabic representations, presumed to be in localized to the posterior superior temporal lobe, are higher-level sensory goals. A somatosensory target in speech production, for example, would code for the cyclical opening and closing of the vocal tract, with

---

[1] Here the psycholinguistic concept of lemma refers to the level of lexical selection that is pre-phonological.

the relative positions of articulators as an endpoint. Taking account of coarticulation, the acoustic consequences of these end targets are imprecise, with no exact position of aperture or closure. In running speech there is no one-to-one mapping between acoustic features and the perceptual categorization of speech sounds; therefore, the larger grain size of the syllable (compared to units of representation referenced in other frameworks, such as phonemes or features) would allow for a more consistent acoustic consequence.

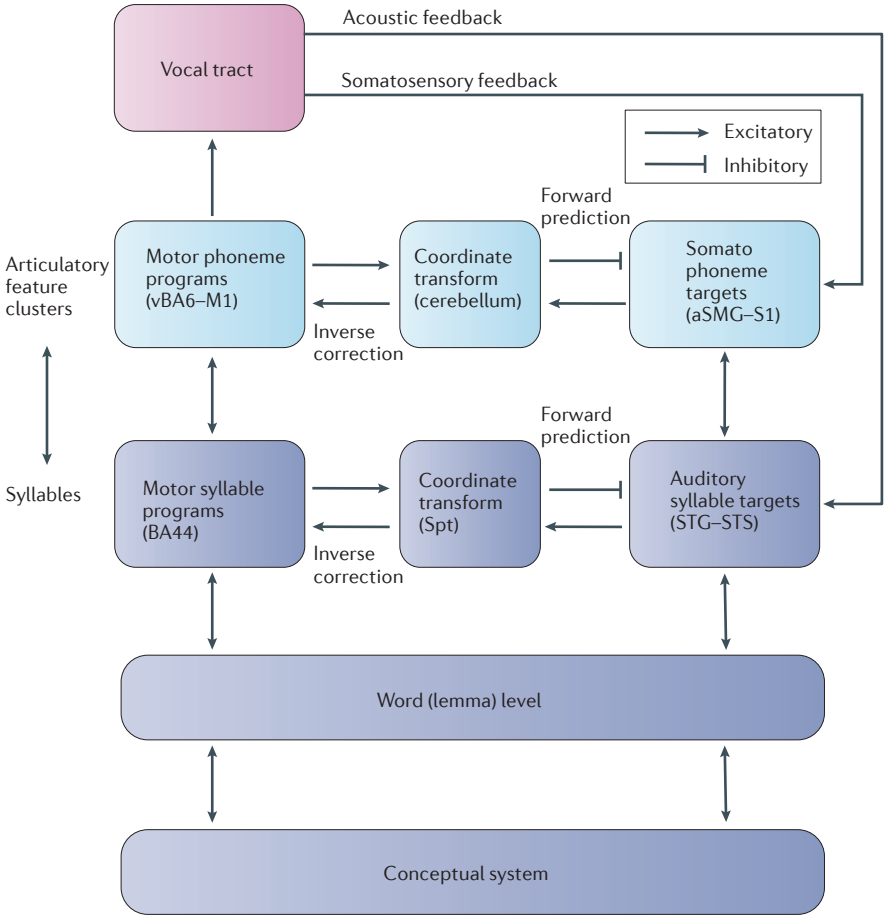


Figure 2. Hierarchical State Feedback Control Model (Hickok, 2012)

2.5.1 The hypothesis of AOS in HSFC

Given the failure of the previously discussed models to provide a cohesive account of AOS, the HSFC provides a viable alternative. Hickok (2012, 2014) loosely discussed how HSFC could account for AOS, noting that the impairment would affect access to the motor-phonological codes, which roughly translate to vocal tract state estimation in the model. Motor planning

necessitates planning over sequences of gestures for coarticulation, rather than a serial, segmental plan. Planning units above the level of the individual phoneme would extend to the syllable. This is consistent with the conclusions of Aichert and Ziegler (2004) that the syllable-sized unit poses a planning and coordination problem in AOS. Hickok provides an explanation of how the model accounts for the close clinical neighbors of AOS: conduction aphasia and dysarthria. In conduction aphasia, the motor system for speech fluency is preserved, as is the phonological system for speech perception, yet the patient struggles with phonemic errors (paraphasias). In HSFC, the core issue in conduction aphasia is internal feedback control, which results from a disconnection between the two systems, where access to information about the targets is lost. Dysarthria, with its consistent and predictable errors owing to muscle weakness, would be a low-level impairment, thus placing AOS as a higher-level impairment with a variable and unpredictable pattern.

2.5.2 Visuomotor processing and the HSFC

Recent neuroimaging work by Venezia and colleagues (2016) tied elements of Hickok's HSFC model of speech production to visuomotor processing. In this study, healthy adults were presented with strings of consonant-vowel syllables in three modality conditions: visual, auditory, and audiovisual. While in the fMRI scanner, participants were prompted to perceive and rehearse via covert articulation, perceive and rest, and continuously perceive syllables across the three modality conditions. The authors hypothesized that the visual input in the visual and audiovisual conditions would either increase activation of auditory motor networks due to multisensory integration of audiovisual information, or recruit additional sensorimotor regions for multisensory processing. Analyses revealed that covert rehearsal of syllables in conditions that contained visual articulatory input activated a network that included left inferior frontal gyrus (IFG), insula, caudate nucleus, and right cerebellum. A distinct sensorimotor pathway for visual speech was identified. A network including bilateral posterior superior temporal sulcus (pSTS), left insula, ventral premotor cortex (VPMC), and the inferior parietal lobe were active during rehearsal across modalities, but preferential to visual and audiovisual conditions. Moreover, bilateral precentral sulci, left central sulcus, caudate nucleus, IFG, and medial temporal gyrus (MTG) demonstrated increased rehearsal-related activation given visual

articulatory input, indicating that visual speech representations may access a distinct pathway in the motor system.

Results of this neuroimaging study suggest that a visuomotor speech pathway may send complementary input to the vocal tract articulators. This additional sensorimotor pathway may serve to integrate visual articulatory input with the motor system during speech production. These conclusions support aspects of Hickok's HSFC model. Speech sound representations guide speech production, thus demonstrating a link between auditory information and the motor system. Internal feedback circuits that engage phonological representations to guide online speech production in real time are tuned by this auditory input. Visual input may similarly engage high-level sensory representations of visual speech that guide articulatory commands to the vocal tract. Thus visual articulatory information plays a role in speech production, supported by the HSFC model. This conclusion is aligned with substantial literature supporting the function of audiovisual integration in speech perception and production, reviewed in the subsequent section.

This chapter provided an overview of AOS in the context of evolving frameworks of motor speech disorders and the integration of linguistic and motor processing. In the next chapter, I review the literature on AV integration, the McGurk paradigm, and event related potentials that relate to the study of multisensory speech perception.

Chapter III

AUDIOVISUAL INTEGRATION

3.1 Audiovisual Integration

Our ability to experience the world is shaped by the integration of information from various senses. Perception of speech and language is also a multisensory process in which we integrate information from a talker's face and body with an acoustic speech signal. The notion of visual input aiding auditory perception was first introduced by Sumby and Pollack (1954). Participants were presented with bi-syllabic words in auditory-only and audiovisual conditions with increasing noise, ultimately reaching a speech to noise ratio of -30dB. Intelligibility in noise was determined by the participants' accurate selection of target words from a list. Results demonstrated a widening gap in intelligibility scores between the two conditions as the speech to noise ratio became progressively larger. The speaker's face provided visual input significantly aided the perception of the target words in adverse listening conditions. The findings of this study have direct implications for how speech is perceived in real world contexts, where background noise is the norm rather than the exception.

Another seminal study by Reisberg and colleagues (1987) underscored the role of visual articulatory input in speech production. Utilizing a shadow repetition task, the authors demonstrated that audiovisual speech input facilitated production during complex speech repetition tasks. Participants were asked to shadow, or repeat upon listening, speech heard in complex and demanding contexts including accented English, a newly learned foreign language, and semantically and syntactically complex content from literary translations. Results showed that participants' repetition of the content, or tracking, measured in words per minute, was significantly faster given visual articulatory input. Thus, audiovisual (AV) representations of speech are not only perceptually salient but also support processes involved in speech production.

Sumby and Pollack's (1954) findings were influential in spurring several lines of research regarding AV speech perception, including integration of AV information during language development, gain of visual information for elderly and hearing impaired populations in adverse listening conditions, and substantial theoretical work examining the relationship between speech perception and production. Most notably, McGurk and MacDonald (1976) demonstrated that

visual articulatory information alters speech perception. This phenomenon, the McGurk Effect, has become the cornerstone of audiovisual speech perception research in psychology, cognitive science, neuroscience, and linguistics.

3.2 The McGurk Effect

McGurk and MacDonald's influential work demonstrated that visual articulatory information alters speech perception, even in clear contexts free of background noise. In the experiment participants viewed and listened to a speaker saying /ba/ and /ga/ with congruent audio and video, as well as incongruent audio and video with the auditory /ba/ stimulus dubbed over the visual /ga/ stimulus. The percept from the incongruent AV pairing was perceived by participants as /da/, a fusion response with phonetic properties of both stimuli. Thus, both audio and visual information were perceived and together formed a new percept, establishing the primacy of multimodal processing in speech perception.

Manuel, Repp, Studdert-Kennedy, and Lieberman (1983) extended these findings, utilizing the stimuli bilabial /ba/ and labiodental /va/. In this experiment the stimuli were visually close in terms of place of articulation, compared to bilabial /ba/ and velar /ga/ in the original experiment. In this paradigm, auditory /ba/ was dubbed over visual /va/. Participants reported hearing /va/, demonstrating that the visual stimulus overrode the auditory stimulus.

3.2.1 Neurophysiological Studies of the McGurk Effect

Relevant to the current study, the McGurk paradigm has been used in neurophysiological investigations of audiovisual speech perception. The mismatch negativity (MMN), a brain event-related potential, has been a useful tool for investigation of the McGurk phenomenon. The MMN will be described in further detail in the following chapter regarding methods of the current study. In the classic McGurk MMN paradigm a series of congruent AV syllables (standards) are interspersed with a rare incongruent McGurk syllable (deviant), and perception of the less-frequent sounds (referred to as "deviants") results in a negative voltage deflection in recordings of electrical brain activity. This negative deflection is referred to as the MMN. Here, a review of MMN investigations of the McGurk Effect is presented to provide a foundation of its reflection of processes in audiovisual speech perception.

Original work investigating a mismatch component for incongruent AV stimuli utilizing the McGurk Effect was carried out with magnetoencephalography (MEG), a brain imaging method similar to EEG but measuring magnetic field fluctuations rather than voltages associated with neuronal communication. Sams and colleagues (1991) employed the stimuli AV congruent /pa/ (standard) and incongruent auditory /pa/ with visual /ka/ (deviant) in the classic McGurk

paradigm, seeking the mismatch field (MMF), the MEG correlate of the MMN. In this study the MMF was elicited in the auditory cortex in the absence of any acoustic change to the stimulus, providing neurophysiological evidence for the visual influence on speech perception.

Colin et al. (2002) established the existence of a mismatch negativity (MMN) response to the McGurk effect, demonstrating an early interaction between auditory and visual stimuli in the perception of speech. The MMN in this type of task reflects both acoustic and phonetic components in the perception of incongruent stimuli. In an oddball paradigm, Colin et al. recorded cortical potentials in three conditions: auditory alone, visual alone, and audiovisual (McGurk effect), subsequently eliciting an MMN effect around 150 milliseconds (ms) post-auditory onset for the McGurk condition. The results of this study demonstrated that the processing of audiovisual speech information is automatic and pre-cognitive, is neither solely auditory nor visual, and relies on continued phonetic processing. It is unclear how sensory information from different modalities is integrated in the perception of speech. The authors suggested that the identification of an MMN for an AV illusion may potentially provide further insight into how short-term memory accommodates phonetic traces.

Similar to the Colin et al. (2002) study, Saint Amour and colleagues (2007) examined the McGurk-MMN, but controlled for visual reactions to stimuli by subtracting the evoked visual responses from the auditory-visual responses. Utilizing an oddball paradigm, this experiment found three distinct phases of McGurk-MMN activity: MMN response at 174 ms post-stimulus over the left temporal scalp, remaining left-lateralized to about 250 ms; a secondary phase of activity that spread bilaterally to fronto-central scalp with a maximum amplitude at around 290 ms; and ultimately a third phase peaking at about 375 ms with a return to left-lateralized scalp sites. These phases were subsequently sourced to the temporal lobe posterior to the primary auditory cortex bilaterally, right hemisphere activity in the superior temporal gyrus (STG), and two sources of left-hemisphere activity in the transverse gyrus and STG. Saint-Amour et al.'s study further characterizes the MMN elicited by the McGurk effect, demonstrating that the response occurred in the absence of acoustic change and not owing to evoked visual responses. The findings here indicate that previous findings of a very early MMN to the McGurk illusion may have been overestimated, occurring because of a change in visual stimuli rather than an updating of expectancy from an incongruent pairing of audio and visual stimuli.

Mussacchia, Sams, Nichols, and Kraus (2006) challenged previous theories of how audiovisual information is integrated in the perception of speech. Massaro (1998) posited that input from various modalities are processed hierarchically via unisensory streams. These unisensory streams converge later in higher order structures in a feedforward manner. In contrast, Mussachia et al. (2006) hypothesized that the integration of AV information occurs very early and interacts in subcortical structures rather than being processed along unisensory streams that converge in cortical structures. In an EEG experiment comparing brain responses to congruent and incongruent AV stimuli, Mussachia et al. tested this hypothesis in two ways: examining the differences between unimodal acoustic responses and AV responses, and also analyzing the responses to the AV stimuli compared to the unimodal acoustic and unimodal visual responses. The authors suggested that the early brainstem response (~11 ms) to AV stimuli is consistent with early activations of nuclei that are peripheral to the thalamus and cortex. Moreover, they state that this early interaction of AV information indicates that visual information affects the human brainstem response early and that the interaction is a result of processing visual information before acoustic information. These findings are aligned with the literature of crossmodal sensory gating, in which early visual information modulates incoming auditory information, as indexed by a response around 50 ms post-stimulus onset (Lebib, Papo, de Bode, & Baudonniére, 2003). The findings of Musacchia et al. (2006) updated the prevailing model of integration of audiovisual information in the perception of speech, which asserted that information from separate auditory and visual modalities is processed along unisensory streams which ultimately intersect in cortical structures. The model implicated by this experiment indicates that the brainstem does not passively receive modality-specific information, but rather integrates multisensory information quite early in processing.

Kislyuk, Mottonen, and Sams (2008) investigated the visual effect on speech perception using an inverse McGurk paradigm. In this study the incongruent AV syllables were created by changing the acoustic rather than the visual component, which would negate the MMN response to McGurk deviants if the visual stimulus overrides the auditory stimulus. The authors hypothesized that the incongruent visual speech stimulus modifies the neural representation of speech in the auditory cortex because it is processed by the same neural population in the auditory stream. In this case, the inverse McGurk, with alterations to the acoustic rather than visual aspects of the stimulus, would render the deviant auditory component identical to the

standard. Indeed, no MMN was elicited to the deviant McGurk stimuli in the AV condition, indicating that visual processing is influenced by the same neural representations in the auditory cortex.

A recent study by Tse and colleagues (2015) combined EEG with event-related optical signals (EROS), a near infrared light (NIR) technique with a combination of high spatial and temporal resolution, to disambiguate the effects of AV integration from deviance detection in classic McGurk MMN paradigms. The authors theorized that AV integration must occur earlier than deviance detection as unisensory and multisensory processing areas engage before the fused percept reaches conscious perception. They utilized three McGurk oddball conditions: AV, visual only, and AV with multiple congruent stimuli to isolate the effects of deviance detection, AV integration, and visual perceptual processes. They found that activity that was exclusively associated with AV integration occurred early in the inferior frontal gyrus (IFG) from 179-230 ms post-stimulus onset, while general deviance detection occurred in the medial temporal gyrus (MTG) later at 332-383 ms post-stimulus onset. Activity related exclusively to AV integration also occurred in the 332-383 ms time window in the occipital cortex, indicating a later interaction between AV integration and deviance detection. The findings of this study confirm that AV integration occurs early in multimodal areas, which interact with unimodal areas in a top-down manner, before phoneme perception is complete.

3.3 Audiovisual Integration in Aphasia

Few studies have examined audiovisual integration in aphasia. The suggestion that patients with aphasia struggle with audiovisual integration is driven by a hypothesis of impaired phonemic processing in the disorder. Moreover, the discovery that speech discrimination difficulties in aphasia are more pronounced for small phonetic differences, marked by errors that predominantly affect place of articulation and voicing, means that any investigation of speech processing in aphasia lends itself easily to the McGurk manipulation (Blumstein, Baker, & Goodglass, 1977; Blumstein, Cooper, Zurif, & Caramazza, 1977).

Campbell and colleagues (1990) performed the first examination of audiovisual speech processing in four adults with brain damage – two with left hemisphere lesions from stroke and two with right hemisphere lesions resulting in prosopagnosia. Participants were presented with stop consonants, vowels, and words in three modalities – auditory, visual, and an audiovisual condition that included McGurk fusions – and asked to repeat the presented stimuli. Results revealed that one left hemisphere patient, characterized as suffering from "word meaning deafness", demonstrated a visual preference to the McGurk illusion, reporting the visual part of the stimulus rather than the auditory aspect. This patient's repetition performance improved given visual input. In contrast, the other left hemisphere patient, characterized as having pure alexia, did not benefit from lip reading for repetition and did not demonstrate visual categorization of McGurk stimuli. The two patients with right hemisphere prosopagnosia demonstrated normal lip reading abilities but an auditory preference during the McGurk illusion. The authors concluded that their results provide evidence that audiovisual integration is left lateralized and rooted in phonological processing.

Hessler, Jonkers, and Bastiaanse (2012) administered a behavioral syllable identification task in four modality conditions (auditory, visual, congruent audiovisual, and incongruent McGurk) to three native Dutch-speaking aphasia patients. Compared to a group of neurotypical adults, all three of the patients with aphasia demonstrated lower accuracy during syllable identification. Interestingly, in the incongruent McGurk conditions, most of the control subjects demonstrated a preference for a response that corresponded to the visual cues in the stimulus, while the aphasia group did not show a patterned response to McGurk stimuli. Although the aphasia group's reaction times were overall slower than that of the control group, within-group comparisons revealed that the control group demonstrated an increased reaction time to the

incongruent McGurk stimuli while 2/3 patients decreased their reaction times during the congruent audiovisual condition and experienced no change in reaction time to the incongruent McGurk condition. The authors suggested that the control group experienced a slowing in response time due to the "double layer" of processing required to access both unimodal and multimodal information. Conversely, the authors surmise that, due to underlying phonological impairment in the aphasia group, access to unimodal phonological information is damaged, resulting in reliance on multimodal information, as evidenced by improved reaction times to congruent audiovisual stimuli. It is difficult to generalize the results from this study as there were only three patients and all had different aphasia presentations (Wernicke's, anomic, and mixed aphasia).

Baum and Beauchamp (2012) reported on the AV speech processing of a patient with a temporo-parietal lesion. Although the patient's lesion included the posterior portion of the left STS along with significant loss of gray matter in the supramarginal gyrus and the auditory cortex, she demonstrated spared AV integration, and showed behavioral evidence that she was experiencing the McGurk effect. The patient, who was diagnosed with only mild anomic aphasia, demonstrated an increased BOLD response in the right STS compared to healthy adults. Both incongruent McGurk and incongruent non-McGurk stimuli (in which the auditory changes rather than the visual) were presented. The comparison group demonstrated heightened responses to non-McGurk incongruent stimuli. The authors noted that the patient's response was equivalent between incongruent audiovisual conditions. They surmised that her responses were based on attentional modulation of audiovisual information. Further they speculated that the patient's improved speech perception over the 5 years from onset of aphasia (increase from 48% to 87% on auditory lexical decision) may have been due to a greater reliance on multisensory processing during the time that auditory processing was weak. The authors also suggested that the significantly greater right STS activations compared to controls could indicate cortical reorganization of multisensory processing following stroke.

3.3.1 Audiovisual Integration in AOS

While informative, the studies reviewed above provide little information that is generalizable to the population of interest. Three studies to date have specifically examined visual speech perception in adults with comorbid nonfluent aphasia and AOS: one study

exploring audiovisual processing mechanisms, and two studies investigating the role of visual speech input in treatment outcomes. Schmid and Ziegler (2006) theorized that access to auditory and visual streams of speech representations occur at a "supramodal" level, where the two modalities are integrated at a later stage of phonological processing. Further, they wanted to know whether these audiovisual processing abilities are unique to language. Fourteen patients with aphasia, eight of whom had comorbid AOS, and control subjects participated in a matching task requiring processing of speech and nonspeech sounds and gestures across four modality conditions: visual, auditory, bimodal (audiovisual), and crossmodal (auditory presented first then visual). Results indicated that the patients were more impaired at crossmodal matching than unimodal matching, with better performance on nonspeech gestures compared to speech gestures, indicating that the impairment may be unique to linguistic processing. Additionally, the patients did not benefit from the visual information in the bimodal condition compared to the unimodal condition, taken as evidence that the impairment did occur at a later supramodal stage of phonological processing. Presence and severity of AOS predicted crossmodal matching performance for speech sounds while nonverbal apraxia predicted nonspeech crossmodal matching performance, implicating the motor system in perceptual performance with specificity for linguistic and nonlinguistic processing.

Fridriksson et al. (2008) examined therapeutic outcomes with audiovisual input for 10 patients with chronic nonfluent aphasia and AOS. Patients participated in a computerized picture naming treatment in two treatment phases with and without articulatory visual information. Results revealed that participants showed significantly greater improvement in picture naming for trained and novel items following the audiovisual treatment phase. Participants did improve in picture naming for trained targets in the treatment phase without visual articulatory information, but gains were not statistically significant. The authors concluded that while frontal areas typically damaged in nonfluent aphasia and AOS are known to play a role in audiovisual speech perception, engagement of these areas during therapy with a perceptual motor speech task may facilitate improvement in speech and language outcomes.

In a follow-up study, Fridriksson and colleagues (2012) further examined the role of audiovisual input in treatment. The treatment of interest in this study, speech entrainment (SE), consists of 1-minute scripted narratives spoken by a fluent, non-impaired speaker. In this study, 13 patients with nonfluent aphasia and comorbid AOS utilized the scripts in treatment with either

auditory-only (AO) or AV feedback. The outcome measure, number of different words produced from the script, revealed that patients produced twice as many different words given AV feedback compared to AO feedback. This gain in the outcome measure was mediated by AOS severity, such that patients with milder AOS made the greatest gains in number of different words produced. The same patients participated in fMRI scanning in conditions examining SE with AV feedback and spontaneous speech. For both patients and comparison participants, SE with AV feedback was associated with greater activations in bilateral anterior insula, BA 37, BA 47, unilateral left medial temporal gyrus, and the dorsal section of Broca's area. Follow-up diffusion tensor imaging (DTI) analysis revealed ventral connections between these structures, in line with Hickok and Poeppel's (2007) framework, that posits a ventral network encoding the conceptual aspects of speech. The patients underwent fMRI scanning again following a 6-week SE treatment with an AV feedback phase. All patients improved significantly following SE treatment, with skills generalizing to untrained scripts. Treatment-related activation increases were noted in the bilateral cingulate gyrus, precuneus, and right hippocampus; while treatment-related decreases in activation were noted in the posterior-inferior parietal lobe including the supramarginal gyrus.

This comprehensive study revealed several important findings that relate to the HSFC model. When examining activations in response to SE with AV feedback, compared to spontaneous speech, the greatest activation was noted in the lexical retrieval area (left BA 47) and areas responsible for regulating visceral activity, like respiration during speech production, in the anterior insula and BA 47, with ventral connections between them. The authors suggest that this network, consistent with the neuroanatomical framework outlined by the HSFC, may reflect on-line predictions about lexical selection along with respiratory preparation for word or utterance length in speech production. According to Hickok (2012), the motor programs in BA 44 are impaired for patients with AOS, secondary to lesions in this critical area. The work of Fridriksson et al. (2012) suggests that this impairment may be further complicated because of impaired access to articulatory motor programs. SE with AV feedback improved speech production for these patients; hence, Fridriksson et al. suggested that SE with AV feedback may provide a visual gating mechanism that pulls along the motor plans in BA 44 with on-line lexical (BA 47) and respiratory (BA 37) predictions to support fluent speech production.

The literature reviewed in this section indicates that AV speech processing is impaired in AOS, yet additional visual articulatory information facilitates speech perception and production. AOS severity has been shown to be predictive of performance on crossmodal matching tasks, which was poorer in patients compared to their unimodal matching performance (Schmid & Ziegler, 2006). While the results of Schmid and Ziegler (2006) indicated that AV information did not facilitate performance on crossmodal matching, Fridriksson et al. (2008, 2012) demonstrated an increase in speech production measures given additional visual articulatory information in treatment. Improvement in speech production given treatment with AV feedback was associated with neural changes in a network related to lexical retrieval and speech production (Fridriksson et al., 2012). Considering the differences in task demands and experimental paradigms between these studies, one behavioral and the others including treatment and neuroimaging, results appear inconsistent and interpretations are unclear. Further study utilizing EEG methodologies could help to disambiguate the seemingly inconsistent results of these studies and elucidate the time course and processing mechanisms of AV integration in patients with AOS. EEG is an ideal method for examining responses in clinical populations since overt responses are not required as indicators of perceptual processing (Näätänen, Paavilainen, Rinne, & Alho, 2007). The following section will review the use of event related potentials for investigating AV linguistic processing.

3.4 Event Related Potentials for Investigating AV Integration

Electroencephalography (EEG) is a non-invasive technique that monitors the brain's electrical activity with high temporal resolution. EEG indexes brain activation at the scalp by recording the electrical activity generated by large populations of neurons. Although action potentials generated by individual neurons are too small to measure from outside the brain, neuronal populations that fire together create larger summed post-synaptic potentials, which can be measured non-invasively. Event Related Potentials (ERPs) are derived from the continuous EEG recordings offline, and index the averaged time-locked neural responses to the repeated presentation of a cognitive event. Averaging together multiple instances of the same cognitive event enhances signal-to-noise ratios of recordings while removing or minimizing the influence of random unrelated activations (Luck, 2005).

Given the millisecond timing precision of this technique, it is ideal for examining the time-course of extremely rapid processes like auditory discrimination. ERP components are defined by their latency, which is the time required by the brain to evaluate the features of the stimulus, with longer latencies indexing more complex stimuli, different stages of processing, or processing difficulty in clinical populations (Hansenne, 2006). ERPs are also described in terms of the direction (positive or negative) of their amplitude, a measure of the resources utilized in response to the cognitive event. For example, the "P300" indexes a positive voltage deflection in response to a cognitive event that occurred 300 milliseconds previously. Some ERPs are considered exogenous components, reflecting the first neural processing of the physical characteristics of a stimulus. For these early sensory components no attention is required and their amplitude is dependent upon cognitive processing. Other ERPs are endogenous components, requiring conscious participation, and these are considered to index higher cognitive processes like attention and memory. In clinical populations, ERPs allow electrophysiological components representing the onset of dysfunction to be identified and subsequent impaired cognitive stages to be inferred (Rugg & Coles, 1995). The current study examines the MMN, as well as a neighboring endogenous component, the P300, both utilized in studies of clinical populations and AV integration.

3.4.1 MMN

The MisMatch Negativity, or MMN, is an event-related brain potential that has been widely studied in both typical and clinical populations. Generally the MMN is investigated in the auditory modality as an index of auditory change detection. The MMN in EEG and its MEG correlate, MMF (MisMatch Field), are elicited in an oddball paradigm where a series of standard (or repeated) stimuli are interspersed with an unexpected deviant, or oddball stimulus. The presence of the deviant violates sensory expectation, eliciting a mismatch response, manifested as a negative voltage deflection in the ERP occurring relatively early post-stimulus onset (Näätänen, 1990, 1992; Näätänen, Gaillard, Mäntysalo, 1978, 1980; Näätänen & Michie, 1979; Näätänen, Simpson, & Loveless, 1982). The MMN is considered an objective measure of auditory discrimination (Kraus, McGee, Carrell, & Sharma, 1995). Studies employing source localization techniques have revealed two cortical generators of the MMN: a bilateral supratemporal process that generates the supratemporal MMN subcomponent, associated with auditory cortex, and a predominantly right-hemisphere frontal process that generates the frontal MMN subcomponent. The supratemporal component is presumed to be associated with pre-perceptual change detection, whereas the frontal component is presumed to be associated with an involuntary attentional shift related to change detection of the auditory stimulus (Näätänen, et al., 1978; 2007; Giard, Perrin, Pernier, & Bouchet, 1990).

The MMN is typically observed in adults between 170-250 ms post-stimulus onset (Handy, 2005; Luck, 2005). Preceding the MMN are early obligatory sensory responses to the auditory stimuli, the P1-N1-P2 complex. This complex reflects sensory processing via a cognitive matching system that compares the stimulus to a trace in sensory memory (Tremblay, Piskosz, & Souza, 2003). First, the P1 peaks around 50 ms post-stimulus onset and serves as an indicator of suppression of unattended information, reflecting the general level of arousal (Key, Dove, and Maguire, 2005). The N1 component peaks around 100 milliseconds after stimulus onset, while the P200 has a latency of 180 to 200 milliseconds (Key et al., 2005). Subsequent to this complex, the MMN occurs, which is located by subtracting the average response to standard stimuli from the average response to the deviants, yielding a difference wave with a negative peak. This negative peak of the MMN is evaluated by its amplitude and latency, with several factors contributing to size and duration, respectively. Relevant to the current study, the degree of (acoustic or linguistic) discrepancy between standard and deviant stimuli has been shown to

affect the amplitude of the MMN, and when MMN amplitude is increased it also has a shorter peak latency. For example, investigations of the MMN in the linguistic domain have shown that MMN responses to familiar, native-language speech stimuli have a greater amplitude and shorter latency when compared with responses to unfamiliar, non-native language speech stimuli (Cheour et al., 1998; Näätänen, 1997; Winkler et al., 1999.

3.4.1.2 MMN in Speech and Language Disorders

The MMN can be elicited in the absence of attention, thus making it an ideal objective neurophysiological measure of linguistic processing in clinical populations. The MMN is considered an index of auditory sensory memory (Alho, 1995; Näätänen, 1992; Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1989; Näätänen & Winkler, 1999) and can be utilized as a marker of decreased sensory memory duration in clinical populations. Other clinically useful indicators indexed by the MMN include aberrant perception, abnormal attentional control, and cognitive decline (see Näätänen et al., 2012 for a review). Most relevant to the current study is the MMN's ability to index decreased auditory discrimination accuracy, particularly in the linguistic domain. Elicitation of the MMN is contingent upon the central auditory system's formation of a representation of the standard stimulus (Näätänen & Winkler, 1999). Given substantial literature demonstrating that the MMN is sensitive to categorical and cross-linguistic speech perception (Cheour et al., 1998; Dehaene-Lambertz, 1997; Kraus et al., 1995; Näätänen, 1997; Sharma & Dorman, 1999; Winkler et al., 1999), it is an ideal component for measuring auditory discrimination accuracy in speech and language impaired populations.

One study examined auditory MMN responses in children diagnosed with childhood apraxia of speech (CAS). While the etiologies of CAS and AOS are distinct, the two motor speech disorders share common speech output characteristics including inconsistent errors, variable productions, groping behaviors, and distorted phoneme productions. Froud and Khamis-Dakwar (2012) investigated whether phonological overspecification contributed to CAS by comparing neurophysiological responses to phonemic (/ba/, /pa/) and phonetic deviants (/pa/, /p^ha/) in children with CAS and typically developing comparison children aged 5-8 years. the typically developing children demonstrated a typical MMN response to phonemic but not phonetic deviants. In contrast, the CAS group did not demonstrate an MMN response to phonemic deviants, but an immature mismatch response to the phonetic contrast – a positivity related to the

standard stimulus. The response of the CAS group to allophonic differences but not phonemic language-specific differences provides evidence that there could be underlying representational deficits in the developmental motor speech disorder.

The McGurk effect and the MMN have not yet been examined in adults with an acquired motor speech disorder. However, some literature exists regarding the presence of MMN responses in adults with aphasia. To date, MMN investigations in aphasia have sought to determine whether speech-processing deficits exist in this population. Csépe and colleagues (2001) examined MMN responses to pitch deviations in tones and vowel and consonant contrasts in CV syllables in 4 Hungarian patients with aphasia. While 2 patients were diagnosed with Wernicke's aphasia and 2 were diagnosed with Broca's aphasia, findings revealed that MMN attenuation was not specific to aphasia type but rather extent and location of lesion in relation to the temporal lobe. Three of the 4 patients demonstrated MMNs to vowel contrasts and all demonstrated MMNs to pitch deviation in tones. MMN amplitude for speech contrasts was correlated with behavioral sound discrimination accuracy.

Ilvonen and colleagues (2003) examined MMNs in 8 Finnish patients with aphasia over the time course of spontaneous recovery from the acute phase until 6 months post-onset. Patients' diagnostic profiles included Wernicke's aphasia, conduction aphasia, anomic aphasia, global aphasia, and transcortical sensory aphasia. In oddball tasks with duration and frequency deviants in tones the patients showed a significant MMN enhancement at the later time points that correlated with speech comprehension performance on a standardized test battery. Marked amplitude increase was noted in response to sounds presented to the right ear; thus, the authors suggested that the MMN may reflect recovery-related mechanisms in the left hemisphere.

Another study utilizing a mixed group of 8 Finnish patients with various aphasia presentations examined MMN responses to frequency and duration deviants in tones and duration and vowel changes in CV syllables. The MMNs of the patients with aphasia were not significantly different from the control responses in the tone condition. Decreased MMN amplitude was noted for vowel and duration changes in speech sounds for the aphasia group. Behavioral results from a separate discrimination task showed significant group differences for reaction times to duration changes in both speech and nonspeech stimuli while the MMN amplitudes only differed between groups for speech sounds. The authors suggest that the

discrepancy between the behavioral and EEG results in identifying group differences indicates that the tasks tap into different types and stages of processing (Ilvonen et al., 2004).

Auther and colleagues (2010) sought to examine the relationship between auditory comprehension, lesion site, and MMN response to place of articulation changes in CV syllables. 17 English-speaking patients with aphasia with various diagnostic profiles were subgrouped by comprehension performance on the *Western Aphasia Battery* and the *Token Test.* All participants with anterior lesions demonstrated MMNs while those with posterior lesions including the temporal lobe did not. Upon examining the comprehension subgroups, 25% of poor comprehenders showed MMNs compared to 89% of good comprehenders. MMN amplitudes were significantly correlated with auditory comprehension scores reinforcing the use of the MMN as an online index of language processing in this population.

Another study also found a relationship between MMN amplitudes and auditory comprehension. Pettigrew and colleagues (2011) examined MMN responses to real words and nonwords in 6 English-speaking patients with aphasia. As in other studies, the aphasia group demonstrated MMNs similar to those of a control group in a non-speech condition utilizing pitch and duration deviants in tones. Aligned with the response of the control group, the aphasia group demonstrated an MMN enhancement for nonwords presented as deviants compared to real words presented as deviants, although mean amplitudes were lower in response to nonwords. This word-related MMN enhancement was significantly correlated with performance on a behavioral lexical discrimination test and MMN mean amplitudes were significantly correlated with both the aphasia quotient and auditory comprehension scores of the *Western Aphasia Battery.*

The studies cited above utilized small, mixed groups of aphasia presentations and varied cross-linguistically. Overall, while MMN studies in aphasia do not create a clear picture of auditory processing for speech sounds in individuals with nonfluent aphasia or AOS, a trend derived from these studies is that attenuated or absent MMN responses are related to overall language ability and are specific to speech stimuli (Pettigrew et al., 2005).

3.4.2 P300

While the MMN is the primary ERP of interest in this study, an additional component will also be targeted since it is sensitive to linguistic deviants, and appears in the time window immediately following the MMN response. The P300 is a positive peak occurring 300 ms post-stimulus onset. It is a long-lasting component with latencies extending 300-700 ms (Desmedt, 1980; Sutton et al., 1965). Similar to the MMN, the P300 is elicited in oddball paradigms in response to the occurrence of a deviant stimulus interspersed among standard stimuli (Polich & Kok, 1995; Toscano, McMurray, Dennhardt, Luck, 2010). The P300 is also used in studies exploring auditory discrimination, including categorical perception of phonemes. However, the P300 is elicited in active tasks requiring the participant to pay attention and generally to provide some kind of overt physical response to stimuli. P300 is thus considered an index of attention and cognitive processing. Like the MMN, the amplitude of the P300 increases given lower probability of occurrence for the deviant (Duncan-Johnson & Donchin, 1982; Johnson & Donchin, 1982). The degree of difference between the standard and deviant stimuli also impacts the amplitude of the P300, associated with the quantity and degree of neurophysiological activation in response to the stimuli (Polich, 2007). P300 amplitude is inversely proportional to the amount of effort involved in the task, with amplitudes decreasing with increased difficulty. The latency of the P300 extends in the direction of difficulty, with increased difficulty yielding longer P300 peak latencies (Polich, 2007; Polich & Kok, 1995).

The P300 has two subcomponents: the P3a and the P3b. The P3a subcomponent usually follows an MMN response, has fronto-central generators, and a latency of 220-280 ms. It is known to occur when the participant is not required to supply an overt response to deviance detection (Squires et al., 1975). Frontal P3a is implicated in involuntary attention as well as inhibition, and is often the component of interest in studies of attention. While P3a is involved in initial signal evaluation, the P3b is presumed to be related to decisional responses at the end-stage of the cognitive processing stream (Polich & Herbst, 2000). The P3b subcomponent, with parietal generators, is related to memory and context updating (Linden, 2005; Polich, 2003). This component has a longer latency of 280-600 msec. Together, the P300 subcomponents heighten memory operations via the transfer of information from the frontal P3a generators to the parietal P3b generators, and reflect the rapid neural inhibition of ongoing activity (Polich, 2007). Like the MMN, the P300 has also been used to study neurophysiological differences between clinical

populations and healthy controls. Attention and memory dysfunction in psychiatric disorders such as schizophrenia, depression, and chronic alcoholism has been associated with P300 amplitude reduction and latency extension (Bruder et al., 1991; Duncan et al., 1987; Linden, 2005; Porjesz & Begleiter, 2003).

The P300 can be evoked in both auditory and visual oddball paradigms (Fornaryova Key, Dove, & McGuire, 2005). Hessler, Jonkers, Stowe, and Bastiaanse (2012) examined P300 responses in an active oddball McGurk task in order to measure conscious activity related to mismatch detection. The authors compared overt responses to both congruent and incongruent McGurk deviants in an AV oddball task. Stimuli included the standard /pa/, congruent deviants /ta/ and /ka/, and McGurk deviant auditory /pa/ dubbed onto visual /ka/. Although the oddball task was active (meaning that a classification response was required of participants), MMN-type negativities were noted in the earlier time windows. The amplitude of the MMN response to McGurk deviants was significantly more negative than the response to AV standards across the time course (120 ms to 400 ms post-stimulus). The amplitude elicited by the McGurk deviants was also significantly more negative compared to the response to congruent AV deviants in the 200-240 ms time window. In the P300 time window, responses to AV stimuli were compared to auditory-only deviants. The positive amplitude in response to the auditory-only stimuli was significantly larger than that to the AV stimuli, indicating that the additional visual information facilitated deviance detection. In sum, the P300 was found to index facilitation of processing given multimodal information.

3.4.2.1 P300 in Aphasia

Only a few studies to date have examined the P300 response in individuals with aphasia. Musiek and colleagues (1992) examined P300 responses in 20 individuals with brain damage of various etiologies, lesions ranging from unilateral left, unilateral right, and bilateral affecting primarily the temporal or parietal lobe. The purpose of the study was to explore the relationship between P300 responses to a cognitive task (counting rare tones) and lesion site and ear of stimulation. While several participants showed no P300 response, remaining participants grouped as whole showed longer latencies and smaller amplitudes compared to the control group. In patients with unilateral lesions no laterality effects were found for peak latency or amplitude for ipsilateral and contralateral ear stimulation.

Hough and colleagues (2003) examined the utility of electrophysiology in assessing auditory processing (via a divided attention dichotic listening task) in seven individuals with aphasia diagnoses including fluent, nonfluent, and anomic aphasia. Results were discussed individually for each participant with inconclusive results regarding the congruence between electrophysiological and behavioral measures of central auditory processing. Waveforms depicting ERP responses were not included in the study report. Of the nonfluent aphasia patients studied, one with normal hearing showed a clear ear advantage in EEG results, in contrast to the behavioral results. Specific components related to this participant's ear advantage were not discussed. Another participant with nonfluent aphasia had moderate to severe bilateral sensorineural hearing loss and showed no congruence between ERPs and behavioral testing. Finally, a participant with nonfluent aphasia and profound left ear hearing loss showed decreased P300 amplitude for ear competition in dichotic listening, but his results were overall dismissed as inconclusive because of his level of hearing loss.

Relatively more recently, an Italian research group examined P300 responses to rare and frequent tones over a 6-month recovery period in 17 patients with global aphasia. Of these patients, 41% demonstrated a P300 response in the sub-acute recordings. After 6 months the patients who demonstrated P300s early in their recovery showed higher amplitudes and shorter latencies over time. These same patients were noted to evolve into a Broca's aphasia diagnosis. EEG recordings were made every month over the recovery period and the data were characterized as fluctuating and unpredictable. By the end of the study 66% of patients showed a stable P300 response to the task. When examining comprehension scores, the patients with subacute P300 responses showed a correlation with improved comprehension (Nolfe et al., 2006).

A substantial body of literature establishes the MMN as an indicator of audiovisual integration in McGurk paradigms. The MMN also has been useful in examining the auditory perceptual abilities of individuals with aphasia. P300 studies in aphasia to date have investigated cognitive processing utilizing deviance detection of acoustic stimuli but not linguistic stimuli. The studies cited here provide few results that can be generalized to audiovisual integration of linguistic stimuli in patients with nonfluent aphasia and AOS. The following chapter will outline the research questions and hypotheses regarding how the MMN and P300 may be impacted in the multisensory processing of individuals with AOS.

Chapter IV

RESEARCH QUESTIONS & HYPOTHESES

The purpose of this research is to determine whether audiovisual speech perception is impaired in adults with AOS secondary to stroke. Regions damaged in AOS are implicated in AV integration for speech (Matchin et al., 2014; Ojanen et al., 2005). Recent work shows that therapeutic interventions that include AV feedback improve speech fluency and number of different words used by patients with AOS (Fridriksson et al., 2008; Fridriksson et al., 2012), in spite of reportedly impaired access to bimodal representations (Hessler et al., 2011; Schmid & Ziegler, 2006). The proposed study has implications for enhanced understanding of the nature of motor speech impairments, and for the development of more accurate diagnostic criteria and more effective treatments. Examination of involvement of linguistic and higher-level multisensory cognitive processes in AOS would further elucidate the etiology and characterization of the specific features used in differential diagnosis.

4.1 Question 1

Are there AV speech perception differences in individuals with AOS compared to healthy comparison participants, as indexed by the MMN and P300 components?

Hypothesis: Acquired AOS in adults post-stroke is associated with impairments to audiovisual integration for speech. Based on this hypothesis, it is predicted that the MMN and P300 Event-related Potentials will be impacted in this population due to changes in the processing streams associated with speech perception and production, such that AV processing will be reflected by an attention-based P300 rather than a pre-attentional MMN.

Brain regions damaged in AOS are also implicated in AV integration for speech (Matchin et al., 2014; Ojanen et al., 2005). Healthy controls exhibit a McGurk Effect, visual influence on auditory perception, when confronted with incongruent AV stimuli in the McGurk condition (Colin et al., 2002; Colin et al., 2004; McGurk & MacDonald, 1976; Musacchia et al., 2006; Sams et al., 1991; Saint Amour et al., 2007). We will compare MMN responses between congruent and incongruent (McGurk) AV stimuli. We predict that patients with AOS will demonstrate impairment in AV integration, inhibiting the McGurk Effect when presented with incongruent AV stimuli. When confronted with an inverse McGurk AV condition (auditory changes, visual remains constant), comparison participants are not expected to show an MMN, indicating that the MMN response is unique to the visual influence on auditory perception, or integration of audiovisual information. We predict that the AOS group will show similar responses between the McGurk and inverse McGurk conditions. Given that patients with AOS show a processing enhancement given AV feedback (Fridriksson et al., 2009; Fridriksson et al., 2012), an alternative hypothesis is that individuals with AOS will show a P300 response, rather than an automatic MMN response, indicating a compensatory facilitation of processing given additional visual articulatory information (Hessler et al., 2012) in both the McGurk and inverse McGurk conditions. We predict that a separate visual-only control condition will reveal that audiovisual responses are unique to bimodal representations and not due to sensory response to visual change detection, as indicated by absent MMN and P300 responses in this condition for the comparison group. We predict for the AOS group that the visual-only control condition will reveal a potential contribution of heightened sensory responses to visual change via an MMN response.

4.2 Question 2

Is AV speech processing right lateralized in individuals with AOS?

Hypothesis: AV integration for speech processing is left-lateralized in healthy adults. During aphasia recovery, right-hemisphere structures are often recruited to support linguistic processing as a function of neural plasticity. Based on these principles we predict that AV speech processing in individuals with AOS will be greater in the right hemisphere than the left, as indexed by larger amplitudes in the P300 component.

EEG studies of aphasia have shown right-hemisphere recruitment for auditory processing as a function of recovery (Ilvonen et al., 2003). AV speech integration is typically served by the left STS in healthy individuals, with greater BOLD responses for multisensory compared to unisensory stimuli (Calvert, Campbell, Brammer, 2000) and these responses are correlated with the strength of the McGurk Effect within an individual (Nath & Beauchamp, 2012). The literature suggests that multisensory integration may become more important for patients with aphasia while unisensory auditory processing is more effortful, and homologous right hemisphere structures may be recruited during recovery (Baum & Beauchamp, 2012). We predict that for the AOS group, P300 responses to incongruent McGurk and non-McGurk stimuli will be greater in the right hemisphere montage. For comparison participants, we predict the opposite response, with left lateralization of ERP responses to AV incongruency (Saint Amour et al., 2007).

4.3 Question 3

Is there a relationship between AV speech processing differences and observable features / symptoms of AOS?

Hypothesis: Reciprocal connections between sensorimotor and visual systems are recruited for speech perception, and therefore AV speech processing differences are related to speech production features of AOS. Based on this hypothesis, we predict that the timing of the response to incongruent AV information will be correlated with aspects of speech motor performance.

Speech production is guided by sensorimotor phonological representations that engage with visual articulatory information. According to Hickock's (2012) HSFC Model, motor phonological codes, impaired in AOS, have reciprocal connections between the vocal tract and the conceptual-linguistic system via sensorimotor and auditory feedback loops. Additional visual sensorimotor pathways engage with these sensorimotor phonological representations during online speech production in real time (Venezia et al., 2016). Therefore, it is plausible that measurable speech production deficits are related to impaired integration of AV speech information. Fridriksson et al. (2012) showed that AV feedback during treatment improved impaired access to motor phonological programs via a visual gating mechanism that interface with online lexical and respiratory predictions necessary for fluent speech production. Several tasks on the ABA-2 involve taxing the online speech production system, some in conjunction with increasing linguistic complexity. Diadochokinetic Rate (DDK), in which the participant rapidly repeats 1-3 syllable combinations, is a measure of volitional control over the articulators The Increasing Word Length subtest measures the individual's ability to serially sequence the correct number of syllables while simultaneously increasing the linguistic load. The Repeated Trials subtest examines deterioration of speech production on successive repetitions. The Inventory of Articulatory Characteristics of Apraxia (part of the ABA-2) provides a count of speech features across speaking contexts (Dabul, 2000). In the ERP measures, peak latency, the time point at which the maximum or minimum voltage deflection occurs, is an index of the temporal progression of a component (Handy, 2005). We predict that brain-behavior correlations between these ABA-2 measures and EEG peak latencies in the AV condition will show a significant relationship, such that lower performance on speech production tasks is associated with longer peak latency as an index of processing difficulty (Polich, 2007).

Chapter V.

RESEARCH DESIGN & METHODS

The following section describes the research design and methods, including recruitment and participants, standardized testing measures, experiment design, stimuli, and procedure. The data reported in Chapter VII were collected based on the following procedures.

5.1 Materials

5.1.1 Behavioral Assessments

A battery of standardized tests was administered to determine type and severity of aphasia, as well as presence and severity of AOS. Testing was administered by a licensed speech-language pathologist.

The Western Aphasia Battery-Revised

The *Western Aphasia Battery-Revised* (WAB-R; Kertesz, 2006) is a standardized assessment administered to patients aged 18-89 years who have neurological damage due to stroke, traumatic brain injury, or progressive neurological disease. The purpose of the WAB-R is to assess linguistic skills related to aphasia, assess nonlinguistic skills, and determine the type and severity of aphasia. The full version of the WAB-R consists of 8 subtests with 32 tasks. The linguistic skills targeted in the assessment include speech fluency, narrative content, auditory comprehension, repetition and naming, reading, and writing. The nonlinguistic skills targeted in the assessment are drawing, calculation, block design, and praxis.

The WAB-R was standardized on a sample of 150 patients with aphasia (114 due to stroke) and 50 healthy controls (Kertesz & Poole, 1974). Validity of the original *Western Aphasia Battery* was examined against the *Neurosensory Center Comprehensive Examination for Aphasia* (NCCEA; Spreen & Bension, 1968) and scores for all aphasia classifications were correlated at $p \leq .0001$. The WAB also demonstrates high internal consistency, with a Cronbach's alpha coefficient of .91 (Kertesz, 2006). The WAB-R yields several subscores including: Spontaneous Speech, Auditory Verbal Comprehension, Repetition, and Naming and Word Finding. Cutoffs for these scores determine the classification of aphasia subtype and sum together into the Aphasia Quotient (AQ). The AQ is a summary value of the patient's aphasic deficit and is proportional to the severity of aphasia, regardless of type or etiology. Severity ratings for the AQ are as follows: 0-25 very severe, 26-50 severe, 51-75 moderate, 76+ mild. In the current study,

the aphasia classification of interest is nonfluent or Broca's aphasia. Cutoff scores for Broca's aphasia are as follows: Fluency <5, Auditory Verbal Comprehension >3, Repetition <8, Naming and Word Finding <9.

Apraxia Battery for Adults- 2nd Edition

The *Apraxia Battery for Adults- 2nd Edition* (ABA-2; Dabul, 2000) determines the presence and severity of AOS in adolescents and adults. The assessment consists of the following subtests: Diadochokinetic (DDK) Rate, Increasing Word Length, Limb Apraxia and Oral Apraxia, Utterance Time for Polysyllabic Words, and Repeated Trials. The test also includes an Inventory of Articulation Characteristics for observation of characteristic in spontaneous speech, with the intention to aid in treatment planning. Each subtest has cutoff scores to determine severity of presentation. The ABA-2 was standardized on a sample of 40 patients with AOS and 49 control subjects. The test is deemed reliable, with Cronbach's alpha coefficients greater than .90 on all main subtests, and .83 for the Inventory of Characteristics. When examined against the *Porch Index of Communicative Ability* (PICA; Porch, 1981), Neely (1980) found that the original *Apraxia Battery for Adults* (ABA) was a more sensitive assessment tool for identifying AOS. Moreover, in this study and a subsequent examination completed during the development of the ABA-2, the test reliably differentiated between AOS and normal speech, and AOS and other speech disorders (Dabul, 2000).

Severity	DDK rate	Increasing Word Length	Limb Apraxia	Oral Apraxia	Utterance Time for Polysllabic Words	Repeated Trials
None	26+	0-1(A & B)	44-50	44-50	0-15	28-30
Mild	7-25	2-4 (A) 2 (B)	37-43	35-43	16-55	16-27
Moderate	2-6	5-7 (A) 3-5 (B)	25-36	21-34	56-80	5-15
Severe	0-1	8+ (A) 6+ (B)	0-24	0-20	81-100	0-4

Table 2. Cutoff scores for severity levels in ABA-2. Increasing Word Length Subtest has two sections, A & B. (adapted from Dabul, 2000)

ABA-2 subtests of particular interest for this study include DDK rate, Increasing Word Length, Repeated Trials, and Inventory of Characteristics of Apraxia. DDK rate is a measure of volitional control over the articulators. In this task participants repeat syllables (/pʌ/, /tʌ/, /kʌ/) and complex sequences of syllables (/pʌtʌ/, /tʌkʌ/, /pʌtʌkʌ/, /plʌkrʌtʌ/) in rapid succession under a time restriction. The raw score is based on the number of trials a participant can articulate without error in 1 second (/pʌ/, /tʌ/, /kʌ/), 3 seconds ((/pʌtʌ/, /tʌkʌ/), and 5 seconds (/pʌtʌkʌ/, /plʌkrʌtʌ/). The Increasing Word Length Subtest requires participants to repeat words of increasing length and linguistic complexity (e.g. *please, pleasing, pleasingly).* This task is scored based on the deterioration in performance from the shortest to longest word per trial. Like the DDK task, the Repeated Trials subtest also measures deterioration in performance, but over 3 successive repetitions of the same word (e.g. newspaper, motorcycle). The score for Repeated Trials is based on the total amount of change, measured in number of errors, from the first to the last repetition. Finally, the Inventory of Characteristics of Apraxia provides a count of the number of features of apraxia (e.g. inconsistent errors, difficulty with initiation, inability to correct known errors) across speaking contexts of spontaneous speech, reading out loud, and automatic speech.

5.2 EEG Experiment

High density EEG was recorded from participants while they were exposed to experimental stimuli. In the audiovisual McGurk paradigm, participants viewed a standard presentation of congruent auditory and visual information (e.g. articulation of /ba/) interspersed with a deviant presentation of incongruent visual (e.g. articulation of /ga/) dubbed over the original audio stimulus. In EEG experiments utilizing the McGurk effect, the congruent audiovisual presentation of /ba/ is presented repeatedly as the standard stimulus. Infrequent presentation of the McGurk stimulus, the incongruent audio presentation of /ba/ paired with visual presentation of /ga/, generates the MMN (Colin et al., 2002; Colin et al., 2004, McGurk & Macdonald, 1976, Sams et al., 1991). In this paradigm the presented audio is consistently /ba/ and only the visual stimulus changes.

For the present study, an additional audiovisual condition utilizing an inverse McGurk deviant was employed to explore the effects of incongruency as mediated by modality. The inverse McGurk condition (AV inverse, labeled AI) utilizes congruent presentation of auditory and visual /ba/ as a standard stimulus with a change to auditory /ga/ while maintaining visual /ba/ for the deviant stimulus. fMRI studies utilizing the inverse McGurk stimulus report that participants do not fuse responses into a single percept but rather perceive both sounds simultaneously (/b-ga/) or only the auditory aspect of the stimulus (/ga/; Nath & Beauchamp, 2012). EEG studies utilizing this type of contrast report no MMN as the auditory contrast is not influential on the visual aspect of the stimulus (Colin et al., 2002; Kislyuk et al., 2008). In the nonspeech domain, visual deviants in AV pairs elicit MMNs of greater amplitudes than auditory deviants and response to visual deviants predicts overall AV responses (Horvath, Schillberg, & Thomson, 2013). Therefore, the AI condition in the present study would provide a direct contrast with the AV condition, elucidating the influence of modality on incongruent deviance detection. A visual-only (VO) control condition is also necessary to ensure that the derived MMN is due to AV integration processes (visual information changing the auditory percept) rather than responses to change in visual stimulus (Saint Amour et al., 2007). See Table 2 for percept schematic of this dissertation study.

Condition	Standard 80%	Deviant 20%	Deviant Percept
AV	👂 /ba/ + 👁 /ba/	👂 /ba/ + 👁 /ga/	/da/
AI	👂 /ba/ + 👁 /ba/	👂 /ga/ + 👁 /ba/	/ga/ or /b-ga/
VO	👁 /ba/	👁 /ga/	/ga/

Table 3. Schematic of the derived percept for each condition in this dissertation study: audiovisual (AV), audiovisual inverse (AI), and visual-only (VO).

The current study follows a 2x3x2 mixed experimental design, with factors group (AOS vs. Comparison), condition (AV, AI, VO), and electrode locations (left vs. right hemisphere). The dependent variable is peak amplitude across electrodes of interest in the MMN and P300 time windows, 175-300 ms and 300-500 ms respectively. Peak amplitude measures the magnitude of response and is calculated from the baseline to the maximum negative-going (MMN) or positive-going (P300) peak of the ERP waveform (Luck, 2005). Peak amplitude is typically used as the dependent variable in MMN studies involving participants with aphasia (Auther et al., 2010; Csépe et al., 2001; Ilvonen et al., 2003; Ilvonen et al., 2004; Wertz et al, 1998).

Data were analyzed within groups and between conditions as well as across conditions and between groups, within the time windows of interest, to evaluate whether there were significant differences in MMN peak amplitude. Within-subject factors were condition (AV, AI, VO) and electrode location (left, right), and Group was included as the between-subject factor (AOS vs. Comparison). In the final analysis brain-behavior correlations were carried out to identify relationships between characteristics of aphasia and/or AOS and peak amplitude of AV speech processing in the MMN and P300 time windows. Difference waves for the MMN and P300 components were also calculated by subtracting the average response to the standard stimuli from the average response to the deviants. Comparisons between hemispheres and conditions were carried out on the difference waves. The study design is summarized in Table 3 below.

AOS	Comparison
AV left, right	AV left, right
AI left, right	AI left, right
VO left, right	VO left, right

Table 4. Planned comparisons of peak amplitude and latency. Within (vertical) and between (horizontal) group comparisons for audiovisual (AV), audiovisual inverse (AI), and visual only (VO).

5.2.1 Stimuli

The AV stimuli for this experiment were designed to elicit the McGurk MMN. Stimuli were generated by digital recording of a female native American-English speaker saying /ba/ and /ga/. Digital video (Canon Vixia HFR50) and corresponding audio (Blue Mic Yeti Pro, www.bluemic.com) were recorded at a sampling rate of 44.1KHz and a frame rate of 24 images/second, later trimmed for a total duration of 300 ms per token. The places of articulation for /ba/ and /ga/ differ maximally. Since the auditory distinction in this paradigm depends on place of articulation, video segments began in the preparatory articulatory position- closed lips for /ba/ and open lips for /ga/. Speaker was instructed to open mouth minimally. Visual inspection of video segments ensured that jaw aperture was consistent between /ba/ and /ga/ video segments. Video frame was cropped using Apple iMovie to reveal only the speaker's mouth in order to constrain the visual presentation, to avoid eye movement artifacts during EEG recording. The audio tracks were separated from the video and edited in Praat (Boersma & Weenink, 2005) with 50ms rise/fall to avoid click artifacts in the recording, and amplitudes normalized to 70dB. The vowel segment from one /ba/ recording was removed in Praat and the spliced /a/ segment was used with the onsets for both /b/ and /g/ so the only difference in the audio is the consonant segment. The audio track for /ba/ was dubbed over the video tracks of both /ba/ and /ga/, creating congruent (auditory /ba/, visual /ba/) and incongruent McGurk (auditory /ba/, visual /ga/) AV stimuli in Apple iMovie. Onset of the AV stimuli begins with the contrastive articulatory position, closed lips for /ba/ and open mouth for /ga/. The AI condition was similarly created, with an inverse McGurk deviant (auditory /ga/ dubbed onto visual /ba/). VO stimuli consisted of the same 300 ms /ba/ and /ga/ video tracks with audio removed. EEG epochs were segmented to coincide with the onset of the auditory component of stimulus to more

specifically examine the neurophysiological response to the visual influence on auditory perception (Hessler et al., 2013).

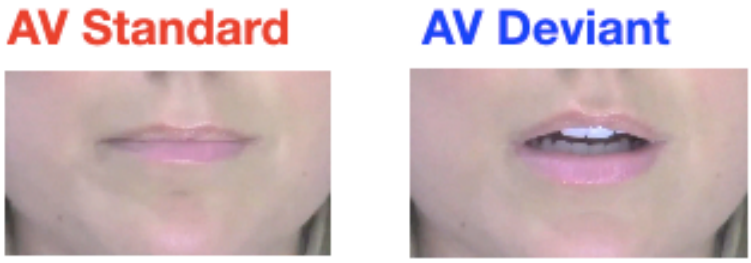


Figure 3. Still shots of video stimuli illustrating frame around the speaker's mouth and onset articulatory position in the AV condition (standard = /ba/, deviant = /ga/).

5.2.2 Oddball Paradigm

Stimuli for the AV, AI, and VO conditions were presented in an oddball paradigm in which the standards were presented for 80% of trials and the deviants for 20% of trials (Luck, 2005). Each condition had 450 total trials, with 360 standards and 90 deviants. Stimuli were presented pseudorandomly in order to ensure that at least two standards came before every deviant and that deviants were not played consecutively. The interstimulus interval (ISI) for all conditions was 600 ms. Presentation of each condition was counterbalanced between participants within each group.

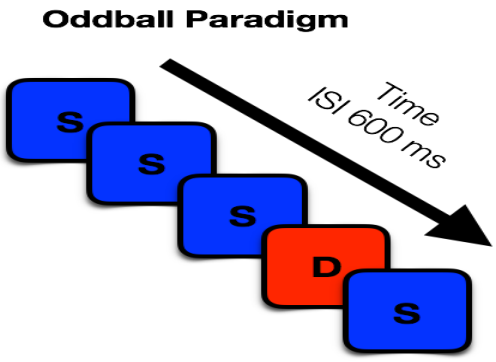


Figure 4. Depiction of oddball paradigm for the current experiment. S= standard, D= deviant. Interstimulus interval (ISI) is 600 ms.

5.3 Participants

A total of 10 participants with aphasia were recruited for this study, from aphasia therapy and support groups at Teachers College and other institutions. Of this group, 4 were excluded based on aphasia type on examination: 3 were found to have anomic aphasia without AOS and 1 was found to have fluent aphasia. The remaining 6 participants had experienced a left hemisphere stroke and all presented with AOS and comorbid nonfluent aphasia. One participant with nonfluent aphasia and AOS was ultimately excluded from the final analysis due to excess noise in the EEG recording. The final analysis for the experimental group therefore includes 5 patients with nonfluent aphasia and AOS (4 males; average age 50.2 years, range 41-68 years). Diagnoses were confirmed via standardized assessment as described in the previous section, and all participant characteristics are summarized in table 3 below. Five healthy comparison subjects, free of neurological disease and speech-language disorders, were recruited and matched for sex, age, and education level (4 males; average age 52.8 years, range 41-69 years). All participants admitted into the study passed a hearing screening at a binaural threshold of 20dBHL across frequencies of 500, 1000, and 2000Hz, and had normal or corrected-to-normal vision. All AOS participants were pre-morbidly right-handed. In the comparison group all participants but one were right-handed. Participants in the AOS group reported no health conditions other than having experienced a left-hemisphere stroke.

Recruitment of participants based on cut-off scores was difficult for the current study for several reasons. The focus of the current study is AOS, which is often comorbid with nonfluent aphasia. Therefore, in order to examine AOS more directly, we investigated AV processing in patients with AOS severity that was equal to or greater than aphasia severity. Additionally, scores on AOS measures are not always proportional to severity, as patients with more severe AOS have less volitional speech and thus demonstrate fewer AOS characteristics overall.

Subject	Age	Sex	DOI	Education	WAB Dx	AQ	ABA Inventory Score
AOS1	41	M	6.5.07	Bachelors	nonfluent/AOS	76.5 mild	8
AOS2	39	F	9.25.13	Masters	AOS	94.6 none	8
AOS3	42	M	5.20.11	Bachelors	nonfluent/AOS	81.2 mild	11
AOS4	64	M	4.15.08	Bachelors	nonfluent/AOS	39.9 severe	14
AOS5	65	M	1.7.07	Doctor of Medicine	nonfluent/AOS	15 very severe	14
COM1	37	M	NA	Masters	NA	98.8	0
COM2	42	F	NA	Bachelors	NA	99.3	0
COM3	48	M	NA	Bachelors	NA	100	0
COM4	68	M	NA	Bachelors	NA	95.6	0
COM5	69	M	NA	Bachelors	NA	100	0

Table 5. Demographic characteristics and select subtest scores for AOS and comparison participants. COM= comparison, DOI = date of incident (stroke, most recent for multiple), WAB Dx = Diagnosis indicated by Western Aphasia Battery, AQ = Aphasia Quotient score on Western Aphasia Battery, ABA = Apraxia Battery for Adults.

5.3.1 Participant descriptions

AOS1: Participant is a 41 year-old male who experienced a left MCA stroke during surgery to correct an anterovenous malformation eight years prior to participation in the study. Participant AOS 1 has a bachelor's degree and worked as a paramedic prior to his stroke. He attends speech therapy regularly and participates frequently in public speaking.

AOS2: Participant is a 39 year-old female who experienced a left MCA stroke 3 years prior to participation in the study. She holds a master's degree and worked as a teacher prior to her stroke. Radiology report from the time of her stroke indicates restricted diffusion in left frontal and temporal lobes. AOS 2 reports that she attends speech therapy infrequently but often participates in research studies.

AOS 3: Participant is a 41 year-old male who experienced a total of 3 left-hemisphere strokes over a ten year period. Patient reported he had a blood-clotting disorder. He has a bachelor's degree and worked in marketing prior to his first stroke. Patient reports that he no longer attends

speech therapy but listens to audio books while reading to support his comprehension and improve his reading skills.

AOS 4: Participant is a 64 year-old male who experienced a left hemisphere stroke 9 years ago. Prior to his stroke participant was a doctor of medicine with specialization in radiology. AOS 4 utilizes an iPad application with stored visual and voice output to support expressive communication.

AOS 5: Participant is a 65 year-old male who experienced a left hemisphere stroke 8 years ago. Prior to his stroke participant earned a bachelor's degree and worked as an air traffic controller. AOS 5 attends speech therapy 3-5 days per week year-round.

5.4 Data Collection

5.4.1 EEG Recording

All EEG recordings took place at the Neurocognition of Language Lab, in the Department of Biobehavioral Studies at Teachers College, Columbia University. The lab employs a 128-electrode high density HydroCel EEG system manufactured by Electrical Geodesics, Inc for EEG recording. The 128 electrodes are arranged in a predictable geodesic position relative to each other in a sensor net. The electrodes are held together by a fine elastomer and contain a silver chloride plated carbon fiber embedded in a plastic substrate. Each electrode has sponge inserts that are soaked in an electrolyte solution of potassium chloride and water before use to ensure optimal conductivity. Sensor nets were selected individually for each participant according to their head circumference. Upon selecting the correctly sized net, additional measurements were made to locate the vertex in order to ensure accurate placement of the net. The participant was then fitted with the appropriate net. Once the net was placed on the participant, he or she was seated in a chair in a sound-attenuated room inside the laboratory. The participant was seated in front of a computer monitor that presents the stimuli. The sensor net was connected to a calibrated amplifier (EGI Net Amps 300 System). Impedance for all electrodes was kept below 40kΩ (Ferree, Luu, Russel & Tucker, 2001). EEG data were recorded using EGI's Netstation (v4.5.4) data acquisition software at a sample rate of 500hz with a sample taken every 2 ms. Following presentation of each condition, electrode impedance was reassessed and electrodes were rehydrated with potassium chloride as needed. The data recording was monitored in real time and bad channels and artifacts were noted and marked so they could be addressed using offline-processing techniques.

5.4.2 Experimental Procedure

Participation in the experiment involved two visits to the lab.

<u>Visit 1:</u>

1. Participants were shown around the lab and familiarized with equipment and procedures by the PI, a speech-language pathologist who has experience working and communicating with individuals with aphasia/AOS. Participants with aphasia/AOS were accompanied by a caregiver if necessary. Questions were actively encouraged and answered during familiarization.

2. Participants were presented with a consent form. The speech-language pathologist allowed the participant and caregiver to read the consent form and also provided the same information in a verbal explanation. Risks were fully explained and any questions were answered before the consent form was signed. Participants were reminded that they could withdraw from the study at any time.
3. Following completion of consent procedures participants participated in administration of standardized tests (ABA-2 & WAB-R).
4. Visit 1 lasted around 45-90 minutes, dependent upon whether the participant needed to go more slowly or take breaks.

Visit 2:

1. Upon determination of appropriateness for the study, the participants were invited back to complete the experimental tasks. On the second lab visit, a consent form (the same as for visit 1) was presented and there were opportunities for questions.
2. Hearing was screened at 500, 1000, 2000, and 4000 Hz within a range of 40-20dB.
3. The head circumference of the participant was measured and the appropriate sensor net was selected. The head was measured further to mark the vertex of the head in order to ensure proper placement of electrodes.
4. The participant was fitted with the appropriate 128-channel HydroCel Geodesic Sensor Net (HCGSN) (Net Amps200, Electric Geodesics Inc., Eugene, OR). Electrodes were referred to the vertex marking made previously on the participant's head.
5. The participant was seated in a chair approximately 80 cm from the computer monitor in a sound attenuated chamber within the Neurocognition of Language Lab. Sounds were presented in free field using a Tannoy OCV 6 full-bandwidth speaker centered 193 cm above the participant's chair. A video camera gave the researcher visual information about the participant during the experiment. The participant was reminded to signal at any time during the experiment if he or she did not wish to continue. The amplifier was checked and calibrated before the net was connected, and impedances (loss of signal between scalp and sensor) were measured. In order to improve impedances the electrodes were adjusted as necessary so that they were in good contact with the participant's scalp.
6. Experimental EEG tasks were presented in random order and counter-balanced across participants. Tasks were presented in short runs of less than 10 min to minimize fatigue

and reduce habituation that interferes will MMN elicitation. Participants were encouraged to take short breaks between runs. The whole EEG experiment lasted less than one hour.

7. Upon completion of experimental tasks, the sensor net was removed and the participant was debriefed. Visit 2 lasted approximately 60 minutes, or longer if the participant needed to go more slowly or take breaks.

5.5 EEG Data Pre-processing

Data pre-processing was carried out via a standard ERP analysis protocol (Handy, 2005; Luck, 2005; Picton et al., 2000) utilizing EGI NetStation (v5.4). The recorded raw EEG data were digitally filtered offline using a 30 Hz LowPass filter. Movement and physiological artifacts (electrocardiogram, electromyography, electrooculography) were removed. Noisy channels were marked as bad and interpolated using spherical spline interpolation based on recorded data from surrounding sensors. Data were re-referenced to an average reference to eliminate the influence of an arbitrary recording reference channel (and also to permit inclusion of the vertex electrode in data analysis). Average referencing instead uses the average of all of the channels (most effectively in a high-density array consisting of more than 64 channels across the scalp; Handy, 2005) to better approximate the ideal zero reference values. The continuous recording was segmented into 750 millisecond epochs, including 100 milliseconds pre-stimulus (the "baseline period") and 650 milliseconds post-stimulus for analysis of the ERP components of interest.

Segments were then averaged together to increase the signal-to-noise ratio and to identify the time-locked event-related responses associated with the onset of the stimuli. EEG epochs were averaged separately for standard trials and deviant trials for each condition, for each individual participant. The averaged waveforms were baseline-corrected to control for drift. Baseline correction procedures involve subtracting the average electrical potential during the 100 ms baseline period from the epoch of interest in order to bring the recording closer to zero, further increasing the signal-to-noise ratio by removing baseline activity that occurred pre-stimulus presentation.

Analysis was constrained to a montage of electrodes of interest in the current study. MMN is maximally recorded at frontocentral electrodes, an area also implicated in AV integration and AOS (Graff-Radford et al., 2014; Hillis et al., 2004; Näätänen, et al., 1978; 2007). Additionally, montages capturing the left and right hemisphere generators of the MMN were included to further examine compensation and recruitment in the AOS group. The left hemisphere montage was composed of electrodes 12, 13, 18, 19, 20, 24; the right hemisphere montage was composed of electrodes 4, 5, 10, 112, 118, 124. Figure 5 shows the regional montages for each area of interest.

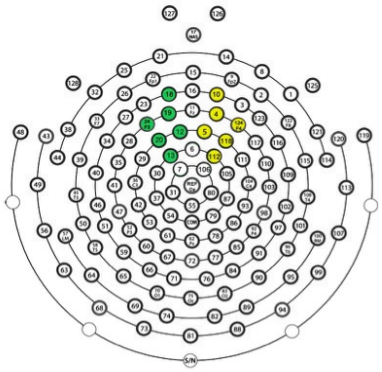


Figure 5. Electrode montage: left (green) and right (yellow) hemisphere montages for MMN and P300 components.

5.1.1 Data Analysis Protocol

Following data pre-processing in NetStation, segmented and averaged EEG data were exported for statistical analysis using SPSS and R Studio (v3.2.2) data analysis packages. A mixed, repeated measures analysis of variance (ANOVA) was conducted on grand-averaged data. The dependent variable was peak amplitude across the time windows of interest (MMN: 175-300 ms; P300: 300-500ms). Data were analyzed within groups and between conditions (AV, AI, VO) as well as between groups in order to evaluate whether there were significant differences in MMN and P300 amplitude. Within subject factors were condition (AV, AI, VO) and location (left, right). Effect sizes for two-way ANOVA are reported as ω^2, considered less biased for small sample sizes. Effect sizes for repeated measures ANOVA are reported as η^2. Both measures represent effect sizes as follows: .01 = small effect, .06 = medium effect, and .14 = large effect (Field, 2013). Significant main effects and interactions were further validated by independent samples t-tests. Pearson correlations were carried out to examine associations between peak latency and behavioral characteristics of AOS. Independent correlations were compared by Fisher-z transformation.

Chapter VI

Processing & Analysis

6.1 Standardized Test Results

The WAB-R and the ABA-2 were administered to determine the type and severity of aphasia and the presence and severity of AOS. Table 5 below summarizes scores on the core subtests for the AOS and comparison groups. Scores for the comparison group were at or near ceiling for all subtests. Administration of the *WAB-R* revealed a diagnosis of nonfluent aphasia for all participants in the AOS group, except AOS 2, whose AQ (94.6) was in the normal range. Participant AOS 1 had an AQ of 76.5 and *ABA-2* scores in the mild range. Participant AOS 2, with AQ in the normal range, also had *ABA-2* scores indicative of mild apraxia. Participant AOS 3's AQ of 81.2 indicates mild aphasia *ABA-2* scores also indicative of mild apraxia. Participant AOS 4's AQ of 39.9 indicates severe aphasia with *ABA-2* scores in the severe apraxia range. Participant AOS 5's AQ of 15 indicates very severe aphasia with *ABA-2* scores also in the severe apraxia range. In summary, 3 of the 5 AOS participants present with none-to mild nonfluent aphasia and mild AOS, and 2 of the 5 AOS participants present with severe nonfluent aphasia and severe AOS.

Participant	SS	Aud	Repetition	Naming	AQ	DDK	Length 1	Length 2	Limb	Oral	Utterance time	Repeat	Inventory
AOS1	14	7.15	7.7	8.5	76.5	18	0	2	49	43	10	30	8
AOS2	19	9.6	9.1	9.6	94.6	14	4	5	50	50	17	27	8
AOS3	18	8.5	6	8.1	81.2	14	9	9	50	50	15	20	11
AOS4	8	5.95	2.2	3.8	39.9	6	3	11	29	29	76	14	13
AOS5	0	4.4	1.3	1.8	15	2	9	NA	50	42	100	2	14
COM1	18	9.8	10	10	95.6	48	0	0	50	50	10	30	0
COM2	20	9.8	10	10	98.8	52	0	0	50	50	10	30	0
COM3	20	10	10	10	100	55	0	0	50	50	10	30	0
COM4	20	10	10	10	100	47	0	0	50	50	10	30	0
COM5	20	10	10	10	100	58	0	0	50	50	10	30	0

Table 6. Assessment results for WAB- R (SS= Spontaneous Speech, Aud = Auditory Comprehension, AQ= Aphasia Quotient) and ABA-2 (DDK= diadochokinetic rate, Length 1 & 2= increasing word length).

6.2 EEG Results

6.2.1 Useable Trials

Following data pre- and post-processing, numbers of useable trials were documented for each participant and mean useable trials were calculated between groups. Although overall the comparison group had numerically more useable trials, there were no significant differences in number of useable trials between groups for each condition. Table 7 below depicts the average number of useable trials in both standard and deviant presentations per group in each condition.

Group	AV Standards	AV Deviants	AI Standards	AI Deviants	VO Standards	VO Deviants
AOS	291 (65)	71 (16)	223 (101)	52 (23)	271 (82)	73 (19)
COM	267 (135)	66 (33)	293 (99)	67 (27)	252 (23)	62 (23)
T-test	t (8) = .36, p = .75	t (8) = .29, p = .79	t (8) = 1.11, p = .37	t (8) = .96, p = .46	t (8) = .34, p = .79	t (8) = .82, p = .43

Table 7. Mean and standard deviation (in parentheses) for number of useable standard and deviant trials for each group in each condition.

6.2.2 EEG Visual Inspection of Waveforms

Prior to analysis, pre-processing procedures described in the previous chapter were implemented. Data were split into separate data sets for responses to standard and responses to deviant stimuli, and segmented with respect to stimulus onset in NetStation. Each standard and deviant data set consisted of values recorded from 128 electrodes sampled every 2 ms, for epochs starting 100 ms prior to stimulus onset and ending 650 ms post-stimulus, for every participant. Data were then reduced to include only the electrodes of interest (i.e., left and right montages). Individual averages were computed by averaging each individual's responses to all standard stimuli and all deviant stimuli, across electrodes within the montage within the epoch of interest, and individual average ERP waveforms were generated (Luck, 2005; Picton et al., 2000). Averaged data were exported to R Studio and SPSS for further analysis. Averaged peak amplitudes within the 175-300 ms time window for MMN and 300-500 ms time window for P300 were examined for the standard and deviant responses of each individual (e.g., Näätänen & Picton, 1987). The MMN response was considered to be present if the average amplitude in response to the deviant stimuli was more negative than the average amplitude in response to the standard stimuli, within the 175-300 ms post-stimulus-onset time window. The P300 response was considered to be present if the average amplitude in response to the deviant stimuli was

more positive than the average amplitude in response to the standard stimuli in the 300-500 ms time window (Hansenne, 2006). Difference waves were calculated by subtracting the average peak amplitude for the standard from that of the deviant for each condition in each time window. Post-hoc analyses were also performed on the difference waves.

Visual inspection of waveforms in the left hemisphere montage revealed a significant MMN response in all 5 comparison participants for the AV condition. In the AOS group, all 5 participants showed a P300 response in the AV condition, with the deviant reflecting more positively than the standard in the 300-500 ms time window. In the AI condition all participants in the AOS group showed P300 responses. In the VO condition 2 comparison participants had a P300 response while 4 out of 5 AOS participants had MMN responses to visual deviants. One AOS participant had a P300 in the VO condition, and another participant had a P300 following the MMN response.

Visual inspection of waveforms in the right hemisphere montage revealed an MMN response in 4 of 5 comparison participants for the AV condition. One comparison participant had a P300 to McGurk deviants and one participant had a P300 that followed the MMN response. In the AOS group, 3 of 5 participants showed a P300 response in the AV condition and 1 participant had a MMN response. In the AI condition all 5 comparison participants had a P300 response, with one participant having an antecedent MMN. In the AOS group 2 participants had MMNs and 2 participants had P300s in the AI condition. In the VO condition all 5 comparison participants had P300 responses, preceded by an MMN for one participant. The AOS group's response to the VO condition in the right hemisphere was less consistent, with 3 participants showing a P300 and one participant showing an MMN. Data were grand-averaged within each group for further analysis, detailed below (see figures below for grand-averaged waveforms for each group and each condition).

6.3 Between-group analyses

ANOVA with factors *group* (AOS vs. COM), *condition* (AV, AI, VO), *status* (standard vs. deviant), and *hemisphere* (left, right) in each time window (MMN, P300) were carried out to evaluate main effects and interactions. In the MMN time window (175-300 ms) no main effects of *group, status, condition,* or *hemisphere* were found, but a significant interaction effect was found for *group x status x condition x hemisphere* ($F(1, 8) = 4.87$, $p = .03$, $\eta^2 = .38$). This indicates a crossover interaction, whereby all variables are affected by all others but no single effect alone reaches significance. Interaction effects were approaching significance for *condition x status* ($F(2, 16) = 3.83$, $p = .06$, $\eta^2 = .324$) and *condition x hemisphere* ($F(2, 16) = 4.12$, $p = .06$, $\eta^2 = .34$). In the P300 time window (300-500 ms) significant main effects were observed for *condition* ($F(2, 16) = 5.13$, $p = .03$, $\eta^2 = .39$), *status* ($F(1, 8) = 12.74$, $p = .007$, $\eta^2 = .61$), and *hemisphere* ($F(1, 8) = 5.25$, $p = .05$, $\eta^2 = .40$). There were no significant interactions in the P300 time window. Follow-up analyses were performed in each condition to determine the effects of *group* and *status* in each time window; these are reported below.

6.3.1 AV Condition:

In left hemisphere montage two-way ANOVA with dependent variable of peak amplitude *status* (standard vs. deviant) and independent variable *group* (AOS vs COM) revealed a significant main effect of *group* ($F(1, 16) = 5.14$, $p = 0.04$, $\omega^2 = .18$) in the MMN time window (175-300 ms) and in the P300 time window (300-500ms) ($F(1, 16) = 5.54$, $p = 0.03$, $\omega^2 = .19$). Two tailed t-tests confirmed a significant MMN (standard vs. deviant) in the left hemisphere for comparison participants ($t(8) = 2.51$, $p = .04$). Two-tailed t-tests on the difference waves also revealed a significant difference between groups ($t(8) = 3.37$, $p = .01$). In the right hemisphere montage ANOVA results revealed a significant main effect of group ($F(1, 16) = 7.6$, $p = 0.01$, $\omega^2 = .27$) in the MMN time window as well as in the P300 time window ($F(1, 16) = 9.92$, $p = 0.01$, $\omega^2 = .30$). Two-tailed t-tests confirmed a P300 approaching significance for the AOS group ($t(8) = 1.89$, $p = .09$). Difference wave analysis was not significant in the right hemisphere montage.

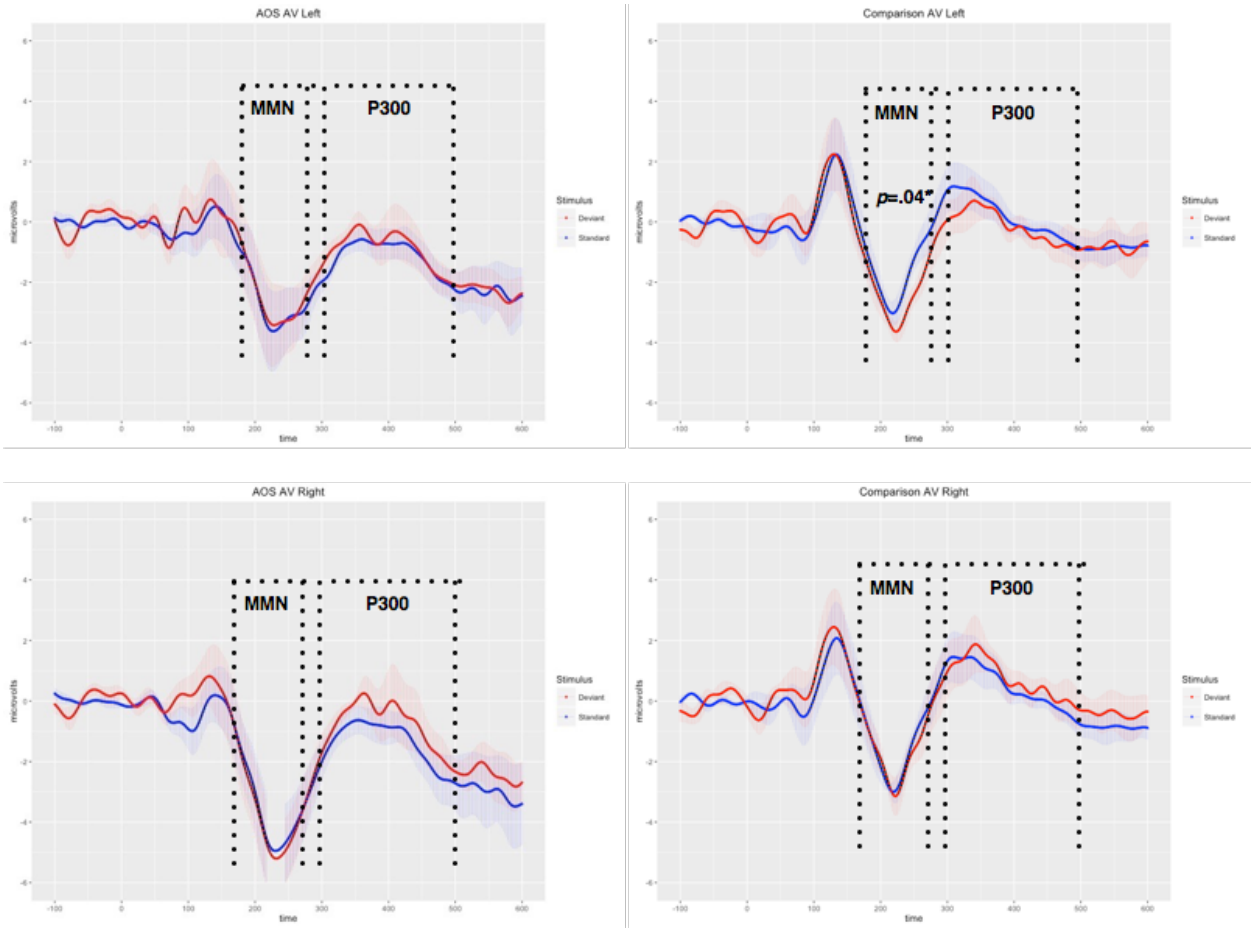


Figure 6. Grand-averaged ERP waveforms for AV condition by group (AOS, Comparison) in left and right hemisphere montages. Shading around lines represents standard error. Standards = blue, deviants = red, measured in microvolts (μv).

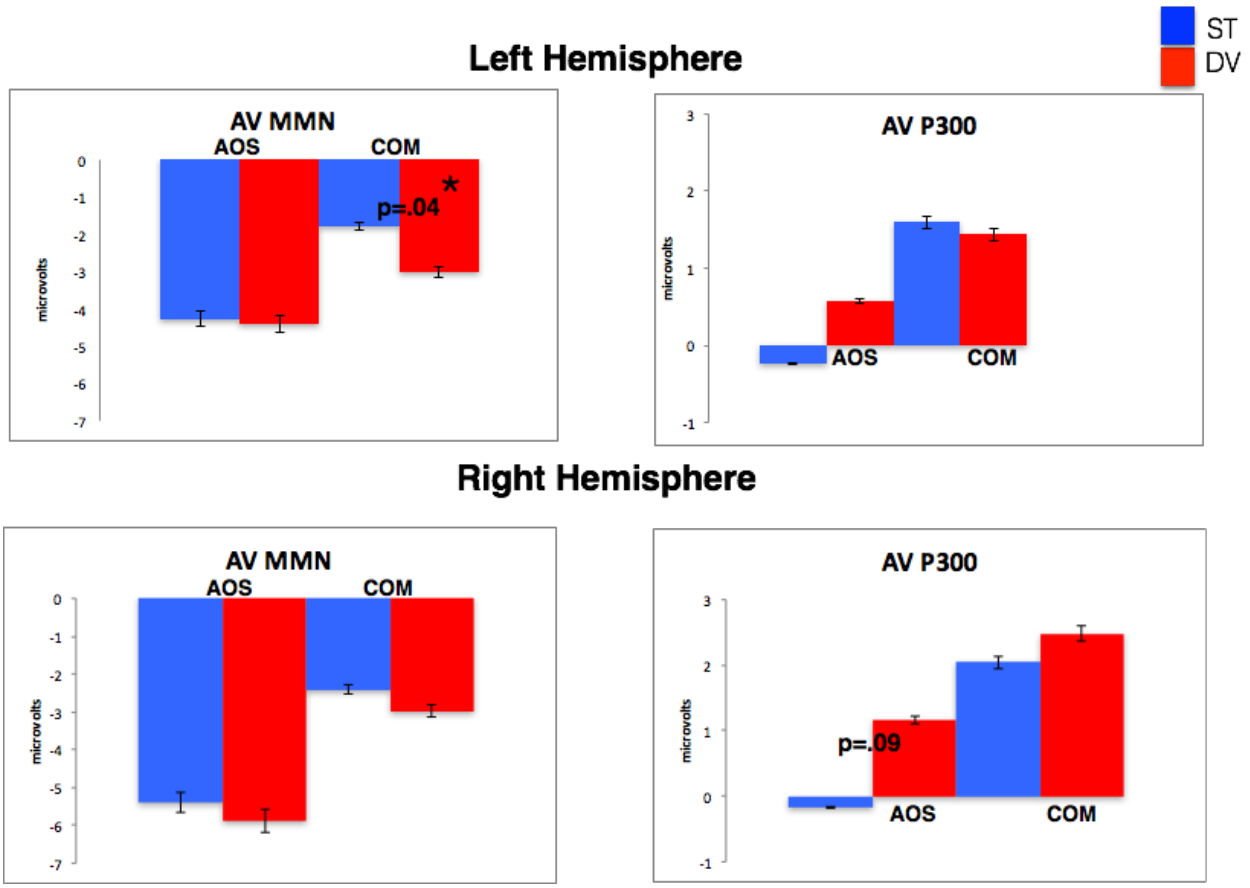


Figure 7. Group mean peak amplitudes for the AV condition in MMN and P300 time windows in the left hemisphere and right hemispheres. Error bras represent standard error of the mean.

	AV Left Hemisphere				AV Right Hemisphere			
	MMN µv	**Latency ms**	**P300 µv**	**Latency ms**	**MMN µv**	**Latency ms**	**P300 µv**	**Latency ms**
AOS ST	-4.27 (2.5)	246 (36.59)	-.23 (1.02)	384 (34.29)	-5.39 (3.31)	248 (31.44)	-.17 (.81)	378 (33.30)
AOS DV	-4.4 (2.7)	246 (27.68)	.57 (1.40)	392 (45.73)	-5.89 (2.86)	234 (17.40)	1.16 (1.34)	384 (37.64)
COM ST	-1.78 (.72)	240 (18.81)	1.59 (1.60)	351 (54.71)	-2.41 (1.28)	244 (14.52)	2.04 (1.42)	350 (55.51)
COM DV	-3.00 (.81)	242 (18.40)	1.43 (.97)	336 (22.33)	-2.98 (1.43)	243 (17.23)	2.48 (1.16)	338 (24.04)

Table 8. Group mean peak amplitude (µv) and latency (ms) with standard deviation in MMN and P300 time windows for the AV condition over left and right hemisphere sensors. ST = in response to standard stimuli, DV = in response to deviant stimuli.

6.3.2 AI Condition:

In left hemisphere montage two-way ANOVA with dependent variable peak amplitude *status* (standard vs. deviant) and independent variable *group* (AOS vs COM) revealed no significant main effects in the MMN time window but a main effect of *group* in the P300 time window (F (1, 16) = 5.43, p = 0.03, ω^2 = .19). In the right hemisphere montage ANOVA results revealed no significant main effects in the MMN time window and a main effect of *group* in the P300 time window (F (1, 16) = 5.34, p = 0.03, ω^2 = .19). Group grand-averaged waveforms are shown below in figure 7. Analyses of the difference waves did not reveal any significant differences between groups in either hemisphere.

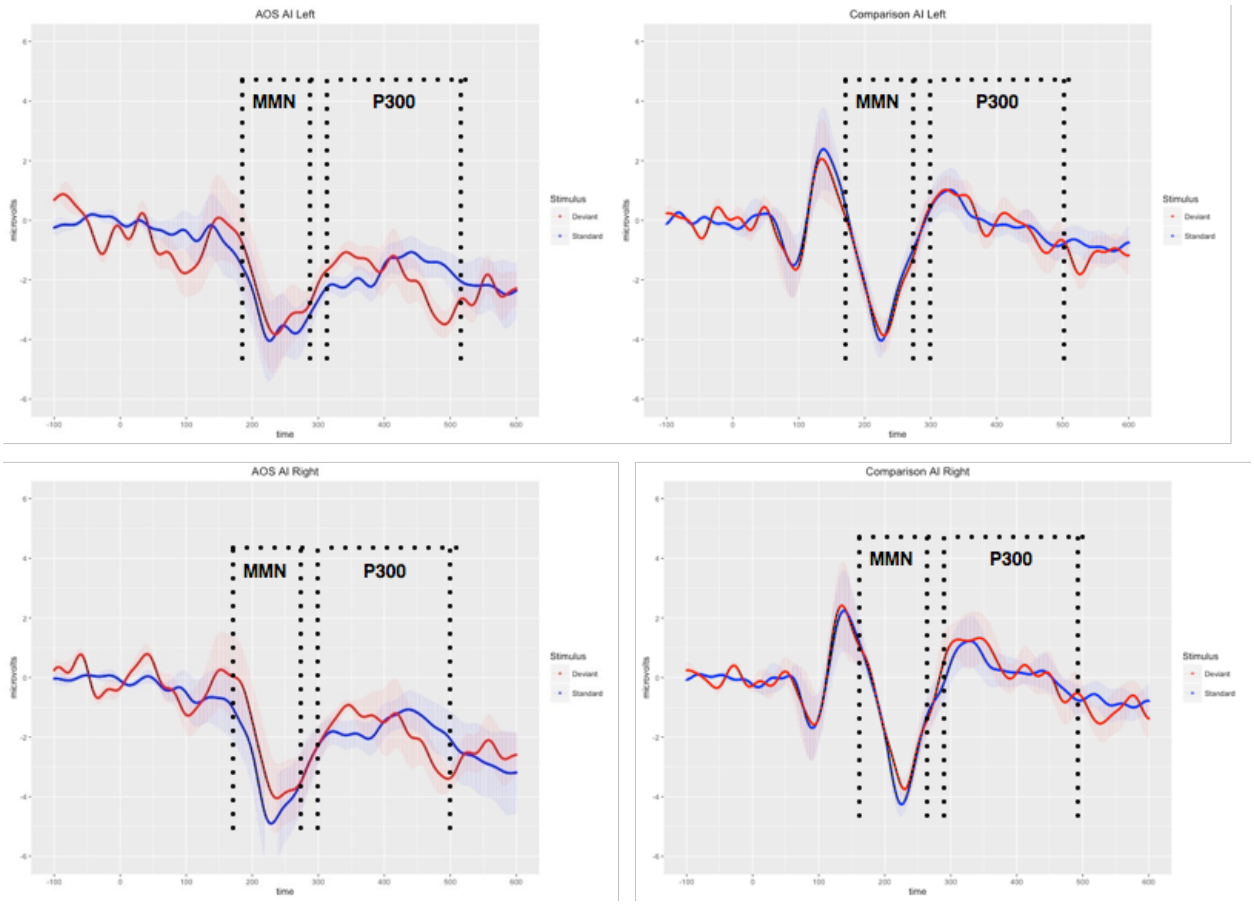


Figure 8. ERP waveforms for AI condition by group (AOS, Comparison) in left and right hemisphere montages. Shading around lines represents standard error. Responses to Standards shown in blue, to Deviants shown in red, measured in microvolts (μv).

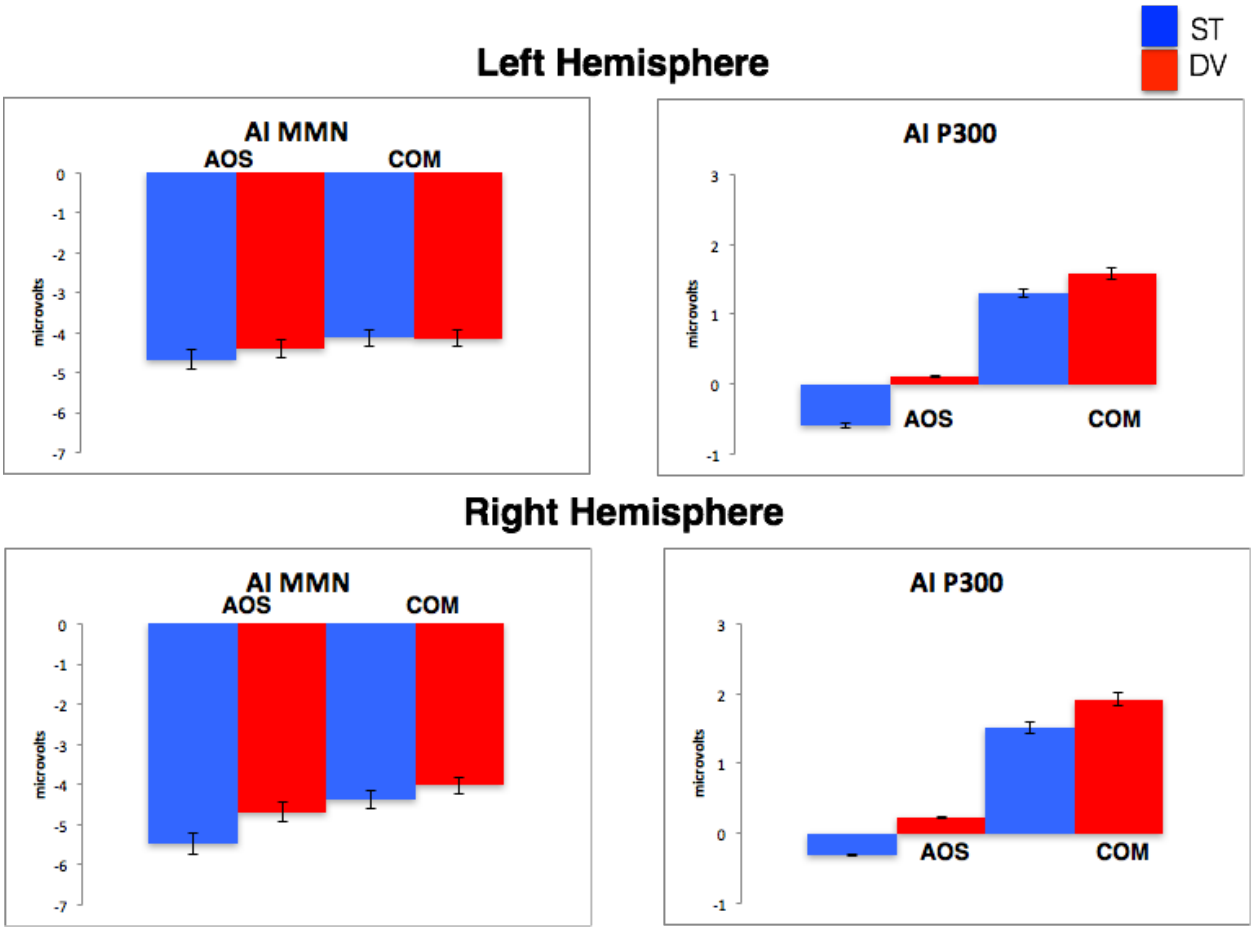


Figure 9. Group mean peak amplitudes for the AI condition in MMN and P300 time windows in the left hemisphere and right hemispheres. Error bars represent standard error of the mean.

	AI Left Hemisphere				**AI Right Hemisphere**			
	MMN µv	**Latency ms**	**P300 µv**	**Latency ms**	**MMN µv**	**Latency ms**	**P300 µv**	**Latency ms**
AOS ST	-4.67 (2.90)	246 (35.62)	-.59 (1.76)	428 (48.34)	-5.39 (3.31)	250 (34.71)	-.17 (.81)	382 (73.22)
AOS DV	-4.41 (4.11)	252 (31.33)	.11 (1.70)	348 (44.84)	-5.89 (2.86)	258 (28.76)	1.36 (1.34)	365 (39.86)
COM ST	-4.13 (1.66)	226 (6.54)	1.31 (1.54)	366 (55.41)	-2.41 (1.28)	228 (7.92)	2.04 (1.42)	360 (54.17)
COM DV	-4.15 (1.0)	228 (9.70)	1.59 (1.45)	372 (52.14)	-2.98 (1.43)	229 (8.56)	2.48 (1.36)	335 (26.82)

Table 9. Group mean peak amplitudes (μv) and latency (ms) with standard deviation in MMN and P300 time windows for the AI condition over left and right hemisphere sensors. ST = in response to standard stimuli, DV = in response to deviant stimuli.

6.3.3 VO Condition:

In left hemisphere montage two-way ANOVA with dependent variable peak amplitude *status* (standard vs. deviant) and independent variable *group* (AOS vs COM) revealed a main effect of *group* in the MMN time window (F (1, 16) = 5.33, p = 0.03, ω^2 = .19) and no significant main effects in the P300 time window. In the right hemisphere montage ANOVA results revealed a significant main effect of *group* in the MMN time window (F (1, 16) = 9, p = 0.01, ω^2 = .31) and no main effects in the P300 time window. Analyses of the difference waves did not reveal any significant differences between groups in either hemisphere.

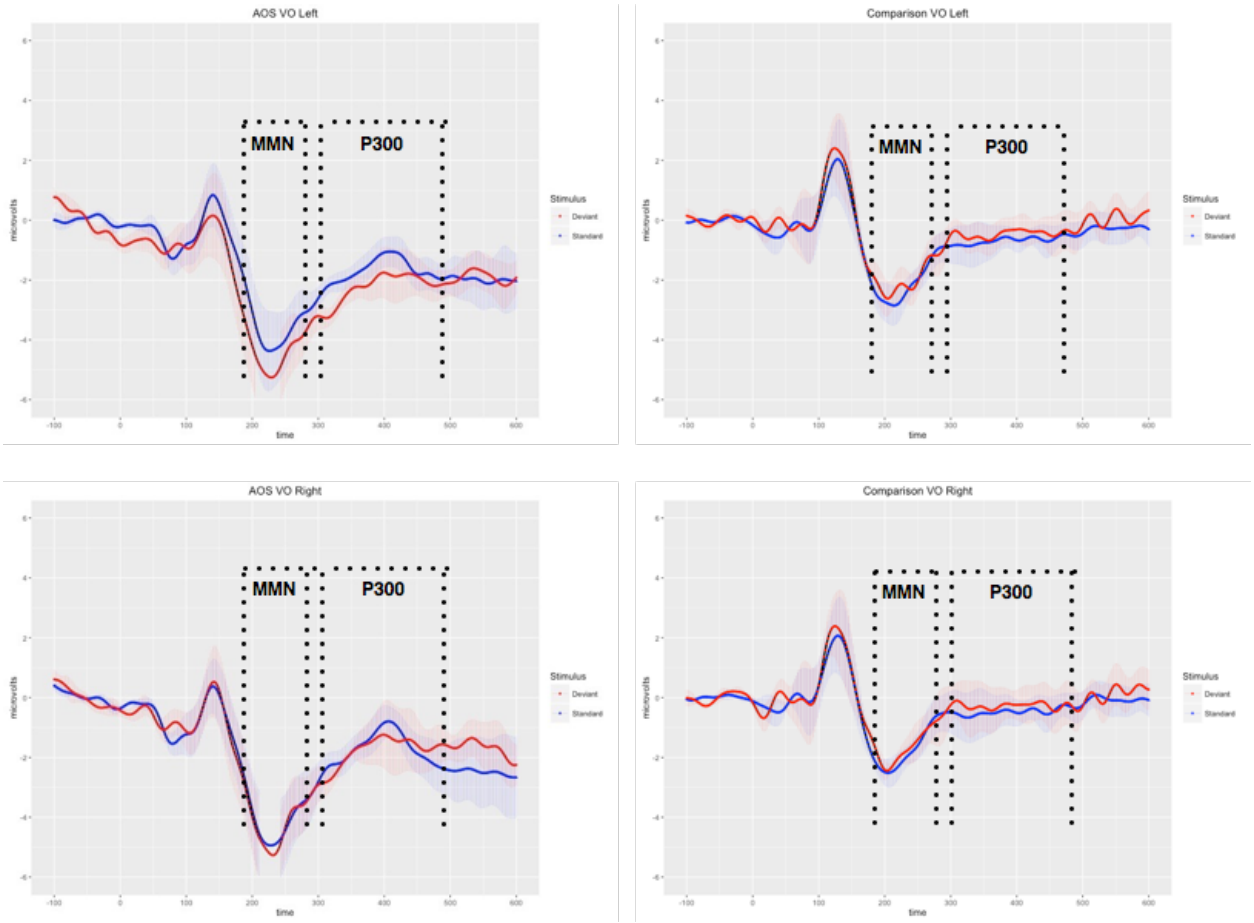


Figure 10. ERP waveforms for VO condition by group (AOS, Comparison) in left and right hemisphere montages. Shading around lines represents standard error. Standards = blue, deviants = red, measured in microvolts (μv).

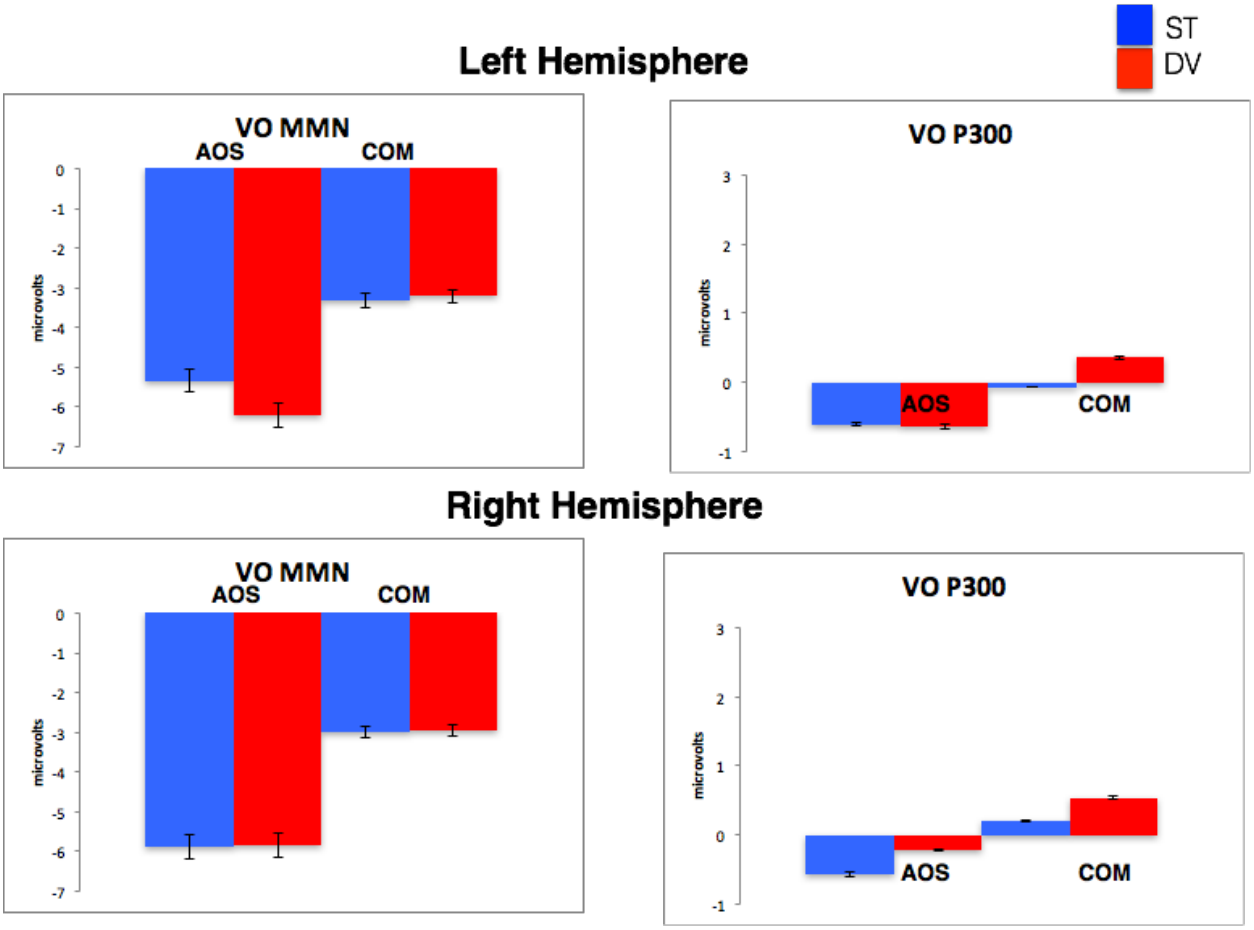


Figure 11. Group mean peak amplitudes for the VO condition in MMN and P300 time windows over left and right hemisphere sensors. Error bars represent standard error of the mean.

	Left Hemisphere VO				Right Hemisphere VO			
	MMN µv	**Latency ms**	**P300 µv**	**Latency ms**	**MMN µv**	**Latency ms**	**P300 µv**	**Latency ms**
AOS ST	-5.33 (3.21)	236 (28.47)	-.61 (1.55)	410 (40.70)	-5.87 (3.0)	238 (25.42)	-.57 (1.10)	411 (17.12)
AOS DV	-6.21 (3.40)	238 (36.92)	-.63 (1.18)	409 (73.63)	-5.83 (2.89)	244 (30.46)	-.22 (1.52)	412 (74.64)
COM ST	-3.31 (1.11)	222 (25.65)	-.07 (.91)	410 (71.93)	-2.98 (.80)	219 (24.23)	.20 (.96)	372 (91.95)
COM DV	-3.2 (.79)	224 (34.18)	.36 (.92)	396 (84.16)	-2.95 (.71)	222 (35.56)	.54 (1.37)	410 (65.75)

Table 10. Group mean peak amplitudes (µv) and latency (ms) with standard deviations in MMN and P300 time windows for the VO condition over left and right hemisphere sensors. ST = in response to standard stimuli, DV = in response to deviant stimuli.

6.4 Within-group analyses

6.4.1 Condition:

2x3 ANOVAs examining the effect of *status* (standard vs. deviant) on *condition* (AV, AI, VO) were carried out within groups for each hemisphere (left, right) and time window (MMN, P300). Difference waves for each condition were also compared.

6.4.1.2 AOS Group:

In the AOS group, left hemisphere analyses revealed no main effects between conditions in neither the MMN nor P300 time windows. In the right hemisphere there was no main effect of condition in either the MMN or the P300 time window. There was a significant effect of *status* (standard vs. deviant) in the P300 time window (F (2, 8) = 9.06, p = 0.04, η^2 = .69). One-way ANOVAs on the difference waves revealed no significant effects between conditions.

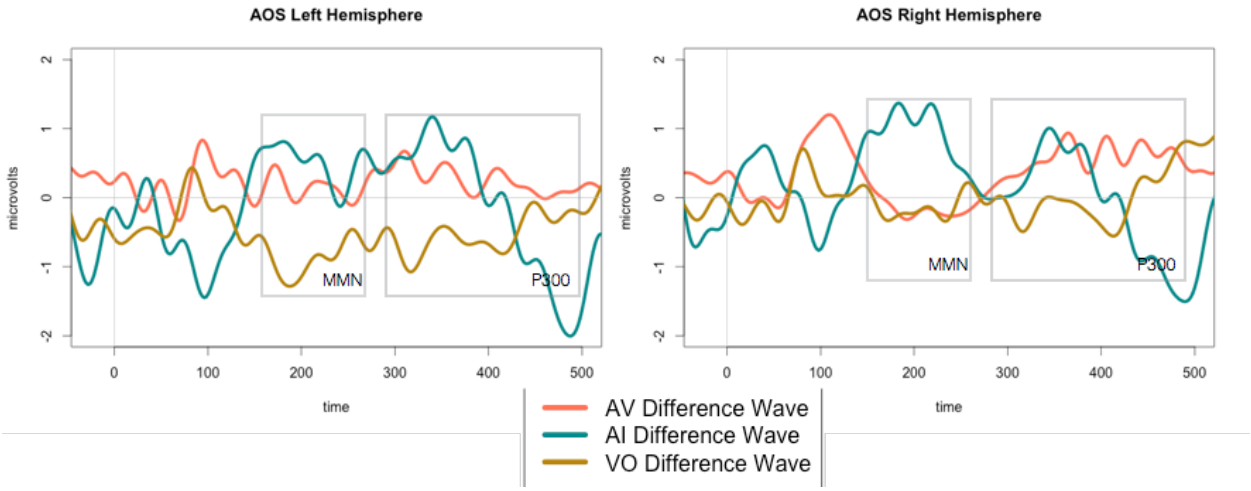


Figure 12. Difference waves for each condition (AV, AI, VO) over left and right hemisphere sensors for the AOS group.

6.4.1.3 Comparison group:

In the comparison group, left hemisphere analyses revealed no significant main effects but a significant condition x status interaction (F (2, 8) = 11.92, p = 0.004, η^2 = .75) in the MMN time window. One-way ANOVA on the MMN difference waves revealed a significant difference between conditions (F (2, 12) = 5.85, p = .02), with significant differences observed between AV and AI (t (3) = 2.44, p = .04), and between AV and VO (t (3) = 4.42, p = .002).

In the P300 time window a significant main effect of condition was found when the non-transformed waveform data were evaluated (F (2, 8) = 6.07, p = .03, η^2 = .60). However, an examination of the difference waves via one-way ANOVA did not reveal significant differences between conditions (F (2, 12) = 1.64, p = .24). Right hemisphere analyses revealed no significant main effects of condition in the MMN time window. In the P300 time window a significant main effect of condition (F (2, 8) = 12.24, p = 0.004, η^2=.75) was found. As before, however, one-way ANOVAs on the difference waves revealed no significant effects between conditions.

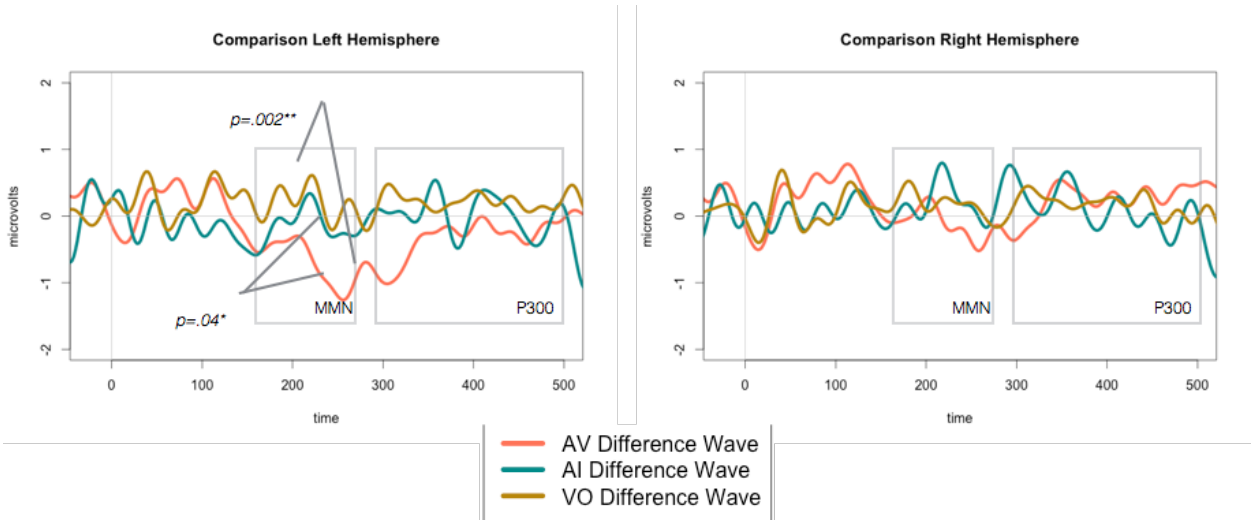


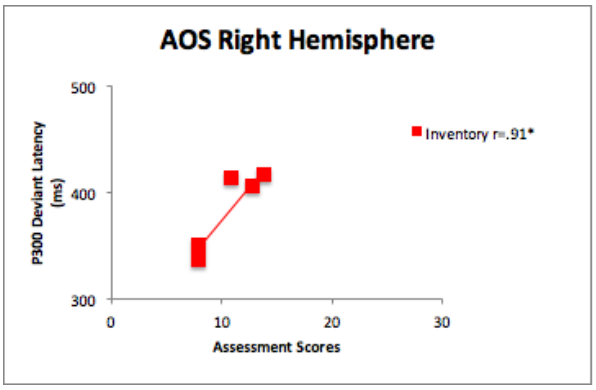
Figure 13. Difference waves for each condition (AV, AI, VO) over left and right hemisphere sensors for the comparison group.

6.5 Brain-behavior analyses

Pearson correlations were conducted to ascertain relationships between peak latency of the components of interest and measures of speech production. Peak latency of the deviant response for the MMN and P300 components for the AV condition were examined in relation to ABA-2 subtest scores for DDK, Increasing Word Length, Inventory of Characteristics of Apraxia, and Repeated Trials. Fisher-z transformations were used to determine significance between two independent correlations.

In the AOS group there were significant correlations between peak latency of the P300 in the right hemisphere montage and Inventory of Characteristics of Apraxia ($r = .91$, $p = .03$) and Repeated Trials ($r = -.84$, $p = .05$). In the left hemisphere montage these correlations were not significant and difference from those of the right hemisphere approaches significance: Inventory of Characteristics of Apraxia ($r = -.33$; $z = 1.57$, $p = .06$) and Repeated Trials ($r = .57$, $z = -1.87$, $p = .06$). No significant correlations were found between DDK or Increasing Word Length with P300 latency in either hemisphere. Correlations between MMN outcome measures and behavioral measures were not computed since this group did not show an MMN response in any of the experimental conditions.

In the control group there were no correlations between subtests of the ABA-2 and peak latency of the MMN or P300 components over either left or right hemisphere sensors. When all participants from both groups were entered into the data pool, significant correlations were found between peak latency of the deviant response of the P300 in the right hemisphere montage with Repeated Trials ($r = -.82$, $p = .004$), Increasing Word Length part 1 ($r = .82$, $p = .004$), Inventory of Characteristics of Apraxia ($r = .78$, $p = .01$), and DDK ($r = -.71$, $p = .02$). Correlations were not significant when behavioral measures and peak latency were compared to left hemisphere sensors. Latency differences between the two hemispheres were significantly correlated with scores on Repeated Phrases ($r = -.11$, $z = 1.96$, $p = .05$), approaching significance for correlation with scores on Increasing Word Length part 1 ($r = .18$, $z = 1.82$, $p = .07$), and non-significant for correlation with scores on Inventory of Characteristics of Apraxia ($r = .56$, $z = .77$, $p = .44$), and DDK ($r = -.56$, $z = .48$, $p = .63$).



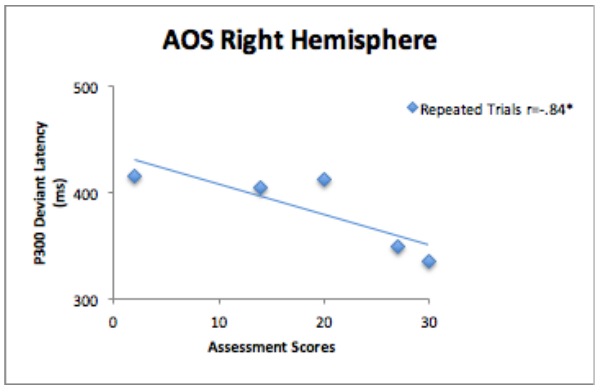
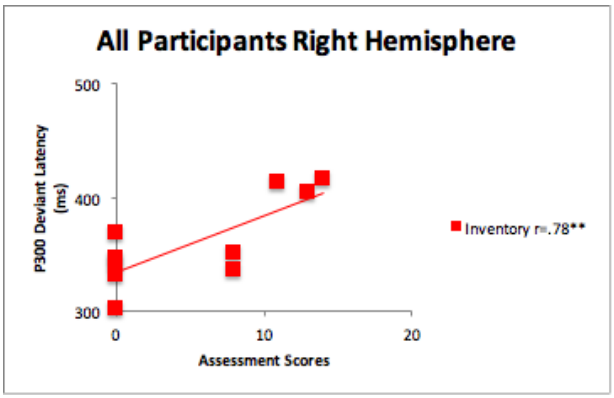
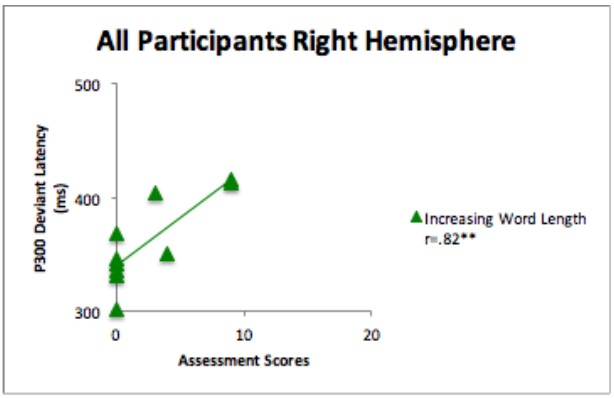
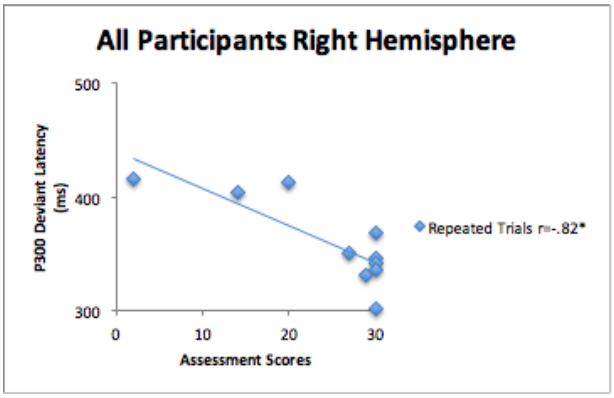


Figure 14. Scatterplots of significant correlations between AV peak latency of the P300 in the right hemisphere and Inventory of Characteristics of Apraxia of speech (left) and Repeated Trials (right) subtests of the ABA-2 for the AOS group.







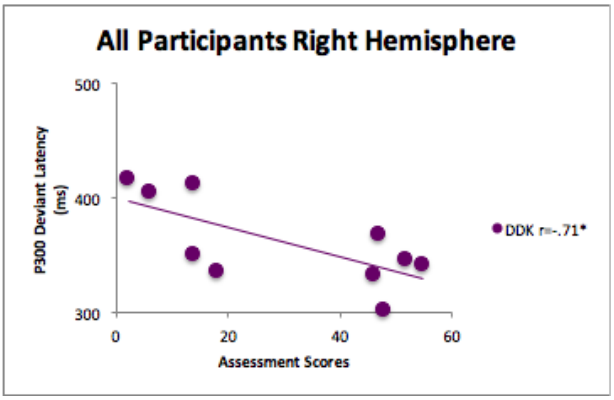


Figure 15. Scatterplots of significant correlations between AV peak latency of the P300 in the right hemisphere and Inventory of Characteristics of Apraxia of speech (top left), Increasing Word Length (top right), Repeated Trials (bottom left), and DDK (bottom right) subtests of the ABA-2 for the all participants combined.

6.6 Results summary

The neurophysiological findings offer some responses to the research questions as follows:

Are there AV speech perception differences in individuals with AOS compared to a healthy comparison group as indexed by the MMN and P300 components?

Is AV speech processing right lateralized in individuals with AOS?

In the AV condition there were significant differences between groups in the MMN and P300 time windows specific to each hemisphere. Over the left hemisphere the comparison group showed a significant MMN, while the AOS group had a P300 response. Over right hemisphere sensors the comparison group's MMN was diminished but the overall differences between group responses were significant. Both groups showed a P300 response over right hemisphere sensors, with significant differences between the groups, likely driven by the larger P300 for the AOS group. In the AI condition there were no significant differences between the groups in the MMN time window over either hemisphere. In both hemispheres the AOS group showed a P300 and group-level differences in responses were significant. The comparison group showed a non-significant P300 response in the AI condition in the right hemisphere. In the VO condition, the AOS group had a large MMN in the left hemisphere, with a significant difference between groups. The comparison group did not show any MMN or P300 for the VO condition in either hemisphere. Distinct profiles were found for each group in each modality and hemisphere. The comparison group showed left-lateralized automatic processing of incongruent McGurk stimuli, indicating AV integration. The AOS group showed a later cognitive-based response indicative of attention-related processing that was not sensitive to differences in type of incongruent AV stimuli (AV vs AI), and earlier automatic responses in the VO condition, suggestive of a reliance on unisensory visual processing.

Hence, in response to the first two research questions, the study showed that there are differences in multisensory speech perception and processing between the two groups as indexed by MMN and P300 ERP components. These differences were apparent in the left-lateralization of the MMN and the larger right hemisphere P300 response of the AOS group.

Are AV speech processing differences related to features of AOS?

Significant correlations were found over the right hemisphere sensors between peak latency of the P300 and Inventory of Characteristics of Apraxia and Repeated Trials. There were no significant correlations over the right hemisphere for the comparison group alone. When both groups were examined together there were significant correlations between P300 peak latency and scores on the Repeated Trials, Increasing Word Length Part 1, Inventory of Characteristics of Apraxia, and DDK sections of the speech assessments. However, MMN latency did not significantly correlate with any of the measures of speech production utilized in this study. The presence of significant correlations when the *n* is higher (as when the groups were collapsed together) raises the index of suspicion for power effects, so the correlations must be interpreted with caution (discussed further below). Nevertheless, this study provides initial evidence that there may be an association between specific behavioral features of AOS and aspects of the brain's engagement in speech processing.

Chapter VII

DISCUSSION

7.1 Discussion

The sensory interpretation of our environment is dependent upon multisensory integration, and thus the alignment of different representational systems. Multisensory integration is a neural response comprised of the interaction of two or more sensory components. This integration occurs when there is a significant difference in neural response between unisensory and multisensory input (Calvert et al., 2000; Stein & Stanford, 2008). Neural integration of multisensory components depends upon both the spatial and the temporal synchrony of the stimuli – they must originate from the same place at the same point in time (Meredith & Stein, 1986, 1996; Meredith et al., 1987; Miller & D'Esposito, 2005). In multisensory integration the unisensory perceptual components are transformed into an integrated product (Calvert, 2001). Speech perception is an inherently multisensory process in which we integrate our visual perception of the movements and expressions of the speaker with the acoustic signal of speech. As exemplified by the McGurk Effect, in speech perception the visual stimulus has the power to override the auditory stimulus, thus changing the percept altogether (McGurk & MacDonald, 1976). This transformed percept has been consistently indexed by the MMN in healthy populations over a variety of manipulations including voiced consonants, voiceless consonants, and place of articulation (Colin et al. 2002; Colin et al., 2004; Hessler et al., 2013; Musacchia et al., 2006; Sams et al, 1991).

Several lines of research suggest that the speech motor system is recruited in mapping acoustic input to phonetic code through articulatory gestures. Experimental paradigms in the field of sensorimotor processing show that forced alterations to the articulators, like perturbation of lip and jaw movement, cause an adjustment to perceived categorical boundaries in acoustic processing. Conversely, downward shifts in vowel formants cause an adjustment to articulatory placement (Lametti et al., 2012; Nasir & Ostry, 2006). Such findings align with the fMRI literature showing that the same speech production motor network that is damaged in AOS is involved in AV integration during speech processing (Ojanen et al., 2005; Skipper, et al., 2005). Thus, in the case of AOS, the impaired speech motor system has consequences for multisensory linguistic processing.

Results revealed AV speech processing differences between AOS patients and comparison participants. The latter demonstrated a significant MMN response to incongruent McGurk stimuli over left hemisphere electrodes. While the response of the comparison group is consistent with substantial literature establishing MMN responses to the McGurk effect, the AOS group demonstrated a different profile: a P300 response to the incongruent McGurk deviants. The AOS group's response to the AI condition was nearly identical to their AV responses. In contrast, the MMN response in the comparison group was unique to the AV condition, with no MMN in the AI condition. The comparison group did not demonstrate a MMN or P300 response to the VO condition, indicating that AV responses were not driven by basic unisensory visual processing. The AOS group did show an MMN response in the VO condition in the left hemisphere, indicating a contribution of evoked visual responses.

The comparison group's MMN response to the McGurk Effect indexes an early, automatic response to incongruent AV information. As the actual auditory stimulus was held constant, the change in the visual stimulus altered acoustic perception for the controls. Saint Amour et al. (2007) used spatial localization techniques to show that this alteration to the percept occurred in the auditory cortex of healthy adults. The AOS group, by contrast, demonstrated a P300 response, associated with cognitive processing and attention. P300 responses are typically evoked in active paradigms in which the participant is required to respond. P300 responses with smaller amplitudes can indicate facilitated processing while larger amplitudes are associated with difficulty during processing (Hansenne, 2006; Polich, 2007). Hessler et al. (2013) used an active AV oddball paradigm, which evoked a small P300 response in healthy adults, interpreted as a facilitation effect for resolving ambiguity. Investigation of the P300 response in a larger sample of AOS participants would provide an opportunity to clarify whether the P300 indexes expedition of processing or a processing cost, through further examination of the impact of AOS on P300 amplitude.

The P300 response of the AOS group indicates that the deviant percept was detected, but this detection of change occurred later and through a different mechanism than that utilized by the comparison group. Behavioral studies of multisensory integration in humans, for both speech and nonspeech stimuli, show that congruency of stimuli decreases reaction times, while incongruent stimuli slow processing (Frens & Opstal, 1995; Sekuler, Sekuler, & Lau, 1997; Stein, Meredith, Huneycutt, & McDade, 1989). Similarly in fMRI studies, incongruent McGurk

stimuli, which are fused via AV integration, elicit a smaller BOLD response than incongruent non-McGurk AV stimuli (Nath & Beachamp, 2012). This facilitative enhancement by congruent multisensory information is supported by the literature on the principle of inverse effectiveness, whereby multisensory enhancement is seen when unimodal effects are weak (Stevenson & James, 2009). Owing to this principle, visual articulatory information aids auditory comprehension in difficult contexts, like speech in noise (Sumby & Pollack, 1954) or in the case of hearing loss (Tye-Murray et al., 2008; Tye-Murray, Sommers, Spehar, Myerson, & Hale, 2010). Neural activation also follows this principle, where congruent audiovisual stimuli evoke a greater response than unisensory auditory or visual stimuli alone (Calvert et al., 1997; Calvert et al., 2000; Stevenson & James, 2009). Another possibility is that the perceptual state of the target is a temporary representation, or object file, that does not necessarily rely on higher level object categorization, thus further dependent upon attention for updating during continuous perception (Kahneman, Treisman, & Gibbs, 1992). Given that the P300 indexes attentional resources, in this case related to AV discrimination, one possible interpretation of the AOS group's response is that it may be associated with the greater resources required to resolve ambiguity. This interpretation aligns with the findings of Fridriksson et al. (2015), that although AV integration is impaired in AOS, additional visual articulatory information is facilitative for both perception and production. Hessler et al. (2013) also interpreted P300 responses to incongruent McGurk stimuli as indicative of a facilitation effect in healthy adults.

The AI condition revealed that the AOS group's responses to traditional McGurk and inverse McGurk stimuli were nearly identical. The comparison group did not demonstrate a McGurk MMN to the inverse deviants in the AI condition, in contrast to the strong MMN elicited by true McGurk deviants. The comparison group's response in the AI condition aligns with the findings of Tse et al. (2015), as AV integration via the visual influence on phoneme perception, occurs early in the frontal lobe. The inverse McGurk deviants in the AI condition manipulate the auditory rather than the visual aspect of the stimulus, which does not generate a fusion response related to AV integration. The AOS group's similar P300 response to traditional and inverse McGurk deviants therefore implies detection of incongruency between modalities that occurred outside the early time window of AV integration.

The VO condition, serving as a control for sensory processing of a change in visual stimulus, did not elicit MMN or P300 responses in the comparison group. The existence of a

purely visual MMN is contested in the literature and is dependent upon specific stimulus features, generally in the nonspeech domain (Fornaryova Key, et al., 2005). The lack of change detection for visual speech in the control group aligns with the body of previous McGurk MMN studies (Sams et al., 1991; Colin et al., 2002, 2004; Saint Amour et al., 2007). In contrast, the AOS group did demonstrate an MMN in the VO condition in the left hemisphere, while their responses in the two bimodal conditions, AV and AI, reflected P300 responses to deviants. An interpretation of these differential effects between conditions may be that patients with AOS attend more heavily to visual speech because phonological processing is reduced. In the VO unimodal condition, change detection happens early in the MMN-time window because there is a reduced linguistic load with no auditory input. Owing to the principle of inverse effectiveness, adults with AOS and nonfluent aphasia may be more sensitive to changes in a visual speech stimulus because they more regularly attend to it in everyday conversational contexts. Baum and Beauchamp (2010) surmised that this phenomenon was responsible for the cortical reorganization of multisensory processing in their case study of a patient with a left temporo-parietal lesion. In the AV and AI bimodal conditions, conflicting auditory and visual stimuli in the incongruent deviants creates a greater linguistic processing load, requiring more effort, and thus a later response to facilitate processing of conflicting information.

The differential ERP responses between groups were also further distinguished by hemisphere. The left-lateralization of the control group's MMNs to the McGurk effect is supported by the results of Saint Amour et al. (2007). Many MMN studies of the McGurk effect in healthy adults utilized a bilateral frontocentral montage. The direct comparison here between left and right showed that the control's MMN was unique to the left hemisphere, as they demonstrated a P300-like response in the right hemisphere. The AOS group showed P300 responses to McGurk deviants, which were numerically but not statistically greater in amplitude over right hemisphere sensors. Similarly, the AOS group showed a P300 in the AI condition to non-McGurk deviants in both hemispheres, while the control group did not. In the VO condition the AOS group had a larger MMN in the left-hemisphere. One interpretation of the AOS group's large MMN in the left hemisphere is similar to that of Baum and Beauchamp (2010): given degraded auditory linguistic processing, visual processing via multisensory integration plays a greater role. This interpretation would also fit with a tendency towards attention-based P300 responses that are trending greater in the right hemisphere, as right hemisphere cortical networks

compensate for the limited contribution of left hemisphere linguistic processing. The analysis here was limited to left and right frontal hemisphere montages, known MMN generators in addition to the area of corresponding to AV integration and the presumed lesions of the patients (Näätänen et al., 1997; Skipper et al., 2007; Tse et al., 2015). Examination of other montages, for example centroparietal electrodes, may provide further clarification of the characterization of the P300 response (Hansenne, 2006; Polich, 2007). Additionally, examination of occipital electrodes may provide further insight regarding the extent of visual processing in AV perception for both AOS and comparison participants.

While these perceptual phenomena revealed by ERPs seem to reflect linguistic processing, the speech production system also plays a role. Significant correlations were found between the peak latency of the deviant in the P300 time window in the right hemisphere and speech production tasks in the AOS group. The significant positive correlation between peak latency of the P300 in the AV condition and the Inventory of Characteristics of Apraxia indicates that the number of diagnostic characteristics noted across speaking contexts, and thus severity, of AOS is related to the timing of the P300 response to incongruent AV information. The significant negative correlation between peak latency of the P300 in the AV condition with the Repeated Trials task suggests that deterioration in speech performance, measured by greater variability and thus a lower performance score upon 3 repetitions of the same word, is also related to the timing of the P300 response to deviant stimuli. These correlations were not present for the control group when analyzed separately. When the data for the AOS and control groups were combined, significant correlations between P300 latency and the behavioral indices from the Repeated Trials and Inventory remained, for the right hemisphere responses to deviant stimuli. A significant negative correlation between DDK scores and peak latency of the right hemisphere P300 was noted for all participants, indicating that the amount of volitional control of the articulators during rapid, timed repetition of syllables and complex syllabic sequences is also related to the timing of deviance detection. A significant correlation between Increasing Word Length and peak latency of the right hemisphere P300 was also present for all participants, such that more errors on a task requiring sequencing of increasingly complex linguistic structures was associated with longer peak latencies.

Although promising, the correlations for all participants combined are interpreted with caution as the control group's responses to Repeated Trials and Increasing Word Length were at

or near ceiling on these speech production tasks, and they showed zero features of apraxia as measured by the Inventory. While Repeated Trials and Inventory of Characteristics of Apraxia were significantly correlated with P300 deviant latency for the AOS group alone, the Increasing Word Length task was only significantly correlated with the ERP measures when additional participants were added to the sample (comparison and AOS groups combined). This correlation may therefore be more indicative of greater power in the sample. Moreover, the nature of these tasks is to expose weaknesses in the motor speech abilities of patients with AOS. Therefore, it would not be expected that the scores on these subtests, whether at or near ceiling, would provide any insight into the AV speech processing of healthy control participants. The DDK task, however, is based on normative data, or the performance of healthy individuals. The distribution of scores for DDK is thus representative of both typical and disordered motor speech abilities. The significant negative correlation between DDK rate and timing of deviance detection implies a relationship between volitional speech motor ability and multisensory speech processing.

The double dissociation of responses between groups highlights two distinct AV processing mechanisms: one that is earlier and more automatic in healthy controls; and one that is later and attention-dependent in the participants with AOS. This differential response seen in individuals with even mild motor speech impairments indicates that higher-level cognitive and linguistic processing deficits may be part of the AOS profile. Difficulties with AV speech integration negate earlier models of speech motor planning or programming as distinct from the linguistic impairment (Aten et al., 1971; Darley et al., 1975) and rather point to an interaction between motor and linguistic processing (Martin, 1974; Hickock 2012; 2014; Laganaro, 2001). In Hickock's (2012) HSFC model, phonological representations are high-level sensorimotor representations. Motor phonological codes, presumably damaged in AOS, correspond to vocal tract estimation. Auditory information from speech sound representations is thus tied to the motor system. Online speech production is tuned by auditory input in real time via internal feedback circuits that engage these sensorimotor (phonological) representations. There is an additional sensorimotor pathway that engages high-level sensory representations of visual speech that guide articulatory commands to the vocal tract (Venezia et al., 2012). According to this model and corresponding neuroimaging data, speech perception and production are accomplished by intricate integration of multiple sensory systems that engage with multifaceted representations of linguistic targets. This advancement in conceptualizing speech representation,

perception, and production align with a potential deficit in AV integration in an assumed motor planning disorder. These preliminary, exploratory results demonstrating differential AV processing profiles between participants with AOS and controls support the notion that AV integration for speech is disrupted in the motor speech disorder.

7.2 Study Limitations and Delimitations

It is essential to note that this study has several limitations. Results of this study were derived from a small sample of only 5 participants in each group, thus underpowering the results. Low statistical power can undermine the results of EEG data (Luck, 2005); however, the relative homogenenity of the AOS responses suggests that the reported findings would likely be consistent with those from a larger participant pool, and provides a foundation for further investigations. The diagnostic profile of the AOS participants varied in severity from mild to severe. One participant had mild AOS without concomitant nonfluent aphasia and one participant had very severe AOS and nonfluent aphasia. Previous MMN studies examining aphasia have also included small numbers of participants, with even greater variation in diagnostic profiles, so despite these limitations the current study is within parameters of pre-existing related literature.

Given the interrelationships between AOS and nonfluent aphasia the results of this study do not fully allow interpretation of the effects of AOS that may be somewhat dissociable from the aphasic linguistic impairment. According to the results of the WAB-R, two patients presented with mild aphasia and one participant did not present with aphasia. These three patients also presented with only mild AOS yet their P300 responses to the bimodal stimuli were consistent with the responses of the participants who presented with severe aphasia and apraxia. Despite the suggestion by these diagnostic profiles that the neurophysiological responses are tied to AOS, along with the significant correlations between the timing of ERP responses and speech motor production performance, the effect of a general linguistic impairment cannot be disambiguated from the speech motor impairment. Indeed, the few existing studies examining AV integration and feedback in this population did not specify the role of a general linguistic impairment on speech motor and AV perceptual abilities (Fridriksson et al., 2009; Fridriksson et al., 2012; Schmid & Ziegler, 2006). Considering current models of speech and language processing such as the HSFC model, the motor speech and linguistic impairment are inter-related and therefore should be considered together.

Another limitation pertaining to this pilot study is the limited lesion data available on the participants. With IRB approval and HIPAA notification, medical history and prior neuroimages (MRI, CT) were requested from participants' medical practitioners, but not produced. Therefore, we cannot confirm the location and extent of the lesions that influence the responses of the

participants. The lack of anatomical information also impacts the analysis of EEG, since activations and connectivity around lesioned brain regions may differ from neurotypical activations in organization, amplitude and latency (e.g. Grefkes & Fink, 2014; Park, Kwon, Kim et al., 2016). There is evidence that density and location of lesion impacts conductivity, influencing source localization algorithms in EEG studies of patients with epilepsy (Brodbeck, Lascano, Spinelli, Seeck, & Michel, 2009). For future studies, obtaining anatomical scans and/or neurological reports could help to determine associations and dissociations between lesion location and behavioral profiles.

While EEG is an ideal method for working with clinical populations because it is non-invasive and does not require overt responses, several challenges remain. Duncan et al. (2009) discuss some obstacles in recording EEG with clinical populations, including poor signal to noise ratio, which reduces data quality. It is noted that several studies of ERPs in patients with aphasia do not publish visual representations (such as waveforms) of participants' responses (Auther et al., 2010; Hough et al., 2003; Musiek et al, 1992; Nolfe et al., 2006). Others publish with smoothed waveforms to ease visual inspection for component identification (e.g., Ilvonen et al., 2003; Ilvonen et al., 2004; Pettigrew et al., 2011). Some studies have not succeeded in using data analysis methods reported with other populations (see Pettigrew et al., 2011). Given the heterogeneity of the experimental groups, some studies present the data qualitatively (Csepe et al., 2001; Musiek et al., 1992, Strouss Hough et al., 2010). Participants in this study had limited usable data due to random noise that cannot be accounted for by environmental factors. The mechanisms by which ERPs are generated in lesioned brains are still not fully understood and require further investigation, especially in conjunction with neuroimaging methods with higher spatial resolution such as MEG and fMRI.

Another limitation of this study is that we cannot determine the faulty mechanism in the AV integration process for the AOS group, only that this process differs from controls and occurs slightly later during speech sound processing. Therefore, we cannot draw any conclusions regarding whether the AOS group experienced AV integration – only that deviance detection occurred around 300ms, significantly later than is seen in neurotypical populations, and that distinct patterns of electrical potentials were observed. Previous studies of AV integration in AOS indicated a severity effect utilizing multisensory information (Fridriksson et al., 2008, 2015; Ziegler, 2012). The small sample size of this study does not allow for in-depth analysis of

severity effects, which may provide additional insights into how AV speech processing relates to speech production. Further investigation of the P300 response in larger sample of AOS participants across severity ranges will clarify whether the P300 indexes expedition of processing or a processing cost by examining the size of the P300 amplitude.

Finally, all participants in this study were in the chronic phase of aphasia. The range of time since onset was 3-9 years in the AOS group. While outside of the acute phase, in which spontaneous recovery is expected (Hillis & Heidler, 2002), varying time between stroke incident and time of testing may have impacted results. Additionally, all five patients reported different amounts of therapy – some currently attend therapy frequently while others rely more on individual compensatory strategies. The intensity of therapy positively impacts recovery (Boghal, Teasell, & Speechley, 2003) and therapeutic intervention may play a role in right hemisphere recruitment for AV speech processing (Baum & Beauchamp, 2012). Therefore, we cannot be certain about the effects of differing types and amounts of treatment on neurophysiological responses, especially when examining right hemisphere recruitment in resolving AV incongruency.

7.3 Conclusions and Future Directions

The findings reported in this study provide preliminary evidence of unique AV processing profiles of participants with AOS and comparison participants. These findings warrant further investigation. Participants with AOS demonstrated the employment of a later, attention-based mechanism when confronted with conflicting AV speech information. Comparison participants, on the other hand, demonstrated an automatic, pre-attentional response that was unique to McGurk deviants or integrated multisensory speech information. These differences were noted across conditions as the AOS group demonstrated P300 responses to inverse McGurk deviants, suggesting that there was no perceptual difference between types of incongruent AV information. The role of visual processing also differed between groups, as the AOS group had a large MMN, perhaps owing to reliance on visual processing in the context of reduced linguistic load. The comparison group did not demonstrate any contributions of visual processing in the visual only condition. The hemispheric differences noted here between groups and between conditions suggest that early processing of AV linguistic information is lateralized to the left hemisphere, whereas the attentional responses to deviance detection are larger over the right hemisphere.

For the AOS group the timing of the P300 responses was correlated with both the number of characteristics of apraxia they demonstrated across speaking contexts as well as performance on motor speech tasks. This suggests that the motor speech system may be involved in the perceptual integration of AV speech information, a possibility that warrants further study.

Interpretation of the results of this study would be better informed with more participants. A comparison group of patients with fluent aphasia and no motor speech involvement may further elucidate the role of the speech motor planning deficits in AV integration for patients with nonfluent aphasia. A more profound understanding of the facilitation effects of AV feedback also has the potential to inform treatment for this population. More advanced data analysis methods, including source analysis, may reveal greater information to aid in the interpretation of ERP responses in individuals with brain lesions.

The findings of differential responses between participants with chronic AOS and healthy controls leads to additional questions regarding whether a disruption to AV speech integration is part of the disease or experience living with a speech impairment. The P300 response may be

interpreted as facilitative, whether for resolving ambiguity or requiring a wider time window for integration. These mechanisms may be compensatory rather than inherent, as suggested by greater reliance on visual information. Further work comparing AV speech processing in individuals in the acute and chronic phases of the disease may provide additional insights.

In conclusion, this study revealed differences in multisensory integration for audiovisual speech information between individuals with a motor speech disorder and a healthy comparison group. These differences signify that a disorder of speech motor planning has the potential to deepen our understanding of the interactions between mechanisms of linguistic representation and those involved in motor speech production.

References

Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkanen, M. (1987). Event-related brain potentials and the perception of a phonetic continuum. *Biological Psychology*, *24*(3), 197–207.

Aichert, I. & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language, 88*, 148-159.

American Speech Language Hearing Association (2015), retrieved from http://www.asha.org/public/speech/disorders/apraxiaadults/

Aten, J. L., Johns, D. F., & Darley, F. L. (1971). Auditory perception of sequenced words in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *14*(1), 131-143.

Basilakos, A., Rorden, C., Bonilha, L., Moser, D., & Fridriksson, J. (2015). Patterns of Poststroke Brain Damage That Predict Speech Production Errors in Apraxia of Speech and Aphasia Dissociate. *Stroke*, *46*(6), 1561-1566.

Besle, J., Fort, A., & Giard, M. H. (2005). Is the auditory sensory memory sensitive to visual information?. *Experimental Brain Research*, *166*(3-4), 337-344.

Blumstein, S. E., Baker, E., & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, *15*(1), 19-30.

Blumstein, S. E., Cooper, W. E., Zurif, E. B., & Caramazza, A. (1977). The perception and production of voice-onset time in aphasia. *Neuropsychologia*, *15*(3), 371-383.

Bhogal, S. K., Teasell, R., & Speechley, M. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke*, *34*(4), 987-993.

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1. 05)[Computer program]. Retrieved December, 2013.

Brodbeck, V., Lascano, A. M., Spinelli, L., Seeck, M., & Michel, C. M. (2009). Accuracy of EEG source imaging of epileptic spikes in patients with large brain lesions. *Clinical Neurophysiology*, *120*(4), 679-685.

Browman, C.P. & Goldstein, L. (1992).Articulatory Phonology: An Overview. *Phonetica, 49* (3–4), 155–180

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., ... & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *science*, *276*(5312), 593-596.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current biology*, *10*(11), 649-657.

Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral cortex*, *11*(12), 1110-1123.

Campbell, R., Garwood, J., Franklin, S., Howard, D., Landis, T., & Regard, M. (1990). Neuropsychological studies of auditory-visual fusion illusions. Four case studies and their implications. *Neuropsychologia*, *28*(8), 787-802.

Center for Disease Control (2015), retrieved from http://www.cdc.gov/stroke/facts.htm

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature neuroscience*, *1*(5), 351-353.

Code, C. (2005). 7 Syllables in the brain: Evidence from brain damage. *Phonological encoding and monitoring in normal and pathological speech*, 119.

Colin, C., Radeau, M., Soquet, A., Demolin, D., Colin, F., & Deltenre, P. (2002). Mismatch negativity evoked by the McGurk–MacDonald effect: A phonetic representation within short-term memory. *Clinical Neurophysiology*, *113*(4), 495-506.

Colin, C., Radeau, M., Soquet, A., & Deltenre, P. (2004). Generalization of the generation of an MMN by illusory McGurk percepts: voiceless consonants. *Clinical Neurophysiology*, *115*(9), 1989-2000.

Croot, K., Ballard, K., Leyton, C. E., & Hodges, J. R. (2012). Apraxia of speech and phonological errors in the diagnosis of nonfluent/agrammatic and logopenic variants of primary progressive aphasia. *Journal of Speech, Language, and Hearing Research*, *55*(5), S1562-S1572.

Csépe, V., Osman-Sági, J., Molnár, M., & Gósy, M. (2001). Impaired speech perception in aphasic patients: event-related potential and neuropsychological assessment. *Neuropsychologia*, *39*(11), 1194-1208.

Dabul, B. (2000). *Apraxia Battery for Adults: Examiner's Manual*. Pro-ed.

Darley F.L. (1969). *The classification of output disturbances in neurogenic communication disorders.* In: American Speech and Hearing Association Annual Conference Chicago IL.

Darley, F., Aronson, A. E., & Brown, J. R. (1975a). Apraxia of speech: Impairment of

motor speech programming. *Motor Speech Disorders*.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975b). *Motor speech disorders* (Vol. 304). Philadelphia: Saunders.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, *8*(4), 919-924.

Desmedt, J. E. (1980). P300 in serial tasks: an essential post-decision closure mechanism. *Progress in brain research*, *54*, 682-686.

Dodd, B. (1977). The role of vision in the perception of speech. *Perception*, *6*(1), 31-40.

Dogil, G., & Mayer, J. (1998). Selective phonological impairment: A case of apraxia of speech. *Phonology*, *15*(02), 143-188.

Dronkers, N. F. (1996). A new brain region for coordinating A. *Nature*, *384*, 14.

Duffy, J. Motor Speech Disorders. St. Louis: Mosby, 1995.

Duffy, J. R., & Josephs, K. A. (2012). The diagnosis and understanding of apraxia of speech: Why including neurodegenerative etiologies may be important. *Journal of Speech, Language, and Hearing Research*, *55*(5), S1518-S1522.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management*. Elsevier Health Sciences.

Duncan-Johnson, C. C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological psychology*, *14*(1-2), 1-52.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., ... & Van Petten, C. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883-1908.

Ferree, T. C., Luu, P., Russell, G. S., & Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology*, *112*(3), 536-544.

Field, A. (2013). Discovering statistics using IBM SPSS Statistics. Fourth Edition. SAGE: London.

Freiherr, J., Lundström, J. N., Habel, U., & Reetz, K. (2013). Multisensory integration mechanisms during aging. *Frontiers in human neuroscience*, *7*.

Frens, M.A. and Van Opstal. (1995) Spatial and temporal factors determine audio-visual

interactions in human saccadic eye movements. *Perception and Psychophysiology, 57,* 802-816.

Fridriksson, J., Moss, J., Davis, B., Baylis, G. C., Bonilha, L., & Rorden, C. (2008). Motor speech perception modulates the cortical language areas. *Neuroimage*, *41*(2), 605-613.

Fridriksson, J., Basilakos, A., Hickok, G., Bonilha, L., & Rorden, C. (2015). Speech entrainment compensates for Broca's area damage. *Cortex*, *69*, 68-75.

Froud, K., & Khamis-Dakwar, R. (2012). Mismatch negativity responses in children with a diagnosis of childhood apraxia of speech (CAS). *American Journal of Speech-Language Pathology*, *21*(4), 302-312.

Gafos, A., & Benus, S. (2006). Dynamics of phonological cognition. *Cognitive Science, 30*, 905–943.

Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia*, *71*, 64-83.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain Generators Implicated in the Processing of Auditory Stimulus Deviance: A Topographic Event-Related Potential Study. *Psychophysiology*, *27*(6), 627-640.

Goldrick, M., Baker, H.R., Murphy, A., Baese-Berk,M. (2011). Interaction and Representational Integration: Evidence from Speech Errors. *Cognition, 121,* 58–72.

Graff-Radford, J., Jones, D. T., Strand, E. A., Rabinstein, A. A., Duffy, J. R., & Josephs, K. A. (2014). The neuroanatomy of pure apraxia of speech in stroke. *Brain and language*, *129*, 43-46.

Green, K. P., & Gerdman, A. (1995). Cross-modal discrepancies in coarticulation and the integration of speech information: The McGurk effect with mismatched vowels. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1409.

Grefkes C. Fink, G. R. (2014). Connectivity-based approaches in stroke and recovery of function. *Lancet Neurol*. 13, 206–216.

Handy, T. (Ed.). (2005). *Event-related potentials: A methods handbook.* Cambridge, Massachusetts: The MIT Press.

Hansenne, M. (2006). Event-related brain potentials in psychopathology: clinical and cognitive perspectives. *Psychologica Belgica*, *46*(1-2).

Hardcastle, W. and Hewlett, N. (2006). *Coarticulation: Theory, Data, and Techniques*.

Cambridge University Press.

Hessler, D., Jonkers, R., & Bastiaanse, R. (2012). Processing of audiovisual stimuli in aphasic and non-brain-damaged listeners. *Aphasiology*, *26*(1), 83-102.

Hessler, D., Jonkers, R., Stowe, L., & Bastiaanse, R. (2013). The whole is more than the sum of its parts–Audiovisual processing of phonemes investigated with ERPs. *Brain and Language*, *124*(3), 213-224.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences, 4*(4), 131_138.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(12), 67_99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402.

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

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*(2), 135-145.

Hickok, G. (2014). Towards an integrated psycholinguistic, neurolinguistic, sensorimotor framework for speech production. *Language, Cognition and Neuroscience*, *29*(1), 52-59.

Hillis, A. E., & Heidler, J. (2002). Mechanisms of early aphasia recovery. *Aphasiology*, *16*(9), 885-895.

Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*, *127*(7), 1479-1487.

Horvath, J. C., Schilberg, L., & Thomson, J. (2013). Does Sight Predominate Sound? Electrophysiological Evidence for Multisensory Mismatch Negativity Correlation. *Neurophysiology*, *45*(5-6), 459-467.

Ilvonen, T. M., Kujala, T., Kiesiläinen, A., Salonen, O., Kozou, H., Pekkonen, E., ... & Näätänen, R. (2003). Auditory Discrimination After Left-Hemisphere Stroke A Mismatch Negativity Follow-Up Study. *Stroke*, *34*(7), 1746-1751.

Ilvonen, T., Kujala, T., Kozou, H., Kiesiläinen, A., Salonen, O., Alku, P., & Näätänen, R.

(2004). The processing of speech and non-speech sounds in aphasic patients as reflected by the mismatch negativity (MMN). *Neuroscience letters*, *366*(3), 235-240.

Itoh M, Sasanuma S, Ushijima T (1979). Velar movements during speech in a patient with apraxia of speech. *Brain and Language*, 7, 227–239.

Johnson, R., & Donchin, E. (1982). Sequential expectancies and decision making in a changing environment: An electrophysiological approach.*Psychophysiology*, *19*(2), 183-200.

Jones, J. A., & Munhall, K. G. (1997). Effects of separating auditory and visual sources on audiovisual integration of speech. *Canadian Acoustics*, *25*(4), 13-19.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(02), 163-178.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive psychology*, *24*(2), 175-219.

Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech.*Journal of Speech, Language, and Hearing Research*, *26*(2), 231-249.

Kertesz, A. (1982). *Western aphasia battery test manual*. Psychological Corp.

Key, A. P. F., Dove, G. O., & Maguire, M. J. (2005). Linking brainwaves to the brain: an ERP primer. *Developmental neuropsychology*, *27*(2), 183-215.

Kraus, N., McGee, T., Carrell, T. D., & Sharma, A. (1995). Neurophysiologic bases of speech discrimination. *Ear and Hearing*, *16*(1), 19-37.

Laganaro, M. (2012). Patterns of impairments in AOS and mechanisms of interaction between phonological and phonetic encoding. *Journal of Speech, Language, and Hearing Research*, *55*(5), S1535-S1543.

Lam, J.M.C. & Wodchis, W. P. (2010) The relationship of 60 disease diagnoses and 15 conditions to preference-based health-related quality of life in Ontario hospital-based long-term care residents. *Medical Care*, 48, 380-387.

Lass, N. (2014). *Speech and Language: Advances in Basic Research and Practice.* Academic Press.

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of aging*, *27*(8), 1155-1163.

Lebib, R., Papo, D., de Bode, S., & Baudonniere, P. M. (2003). Evidence of a visual-to-auditory

cross-modal sensory gating phenomenon as reflected by the human P50 event-related brain potential modulation. *Neuroscience letters*, *341*(3), 185-188.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1-36.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, Mass: MIT Press.

Maas, E., Barlow, J., Robin, D., & Shapiro, L. (2002). Treatment of sound errors in aphasia and apraxia of speech: Effects of phonological complexity. *Aphasiology*, *16*(4-6), 609-622.

Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor programming in apraxia of speech. *Brain and Language*, *106*(2), 107-118.

Maas, E., Gutiérrez, K., & Ballard, K. J. (2014). Phonological encoding in apraxia of speech and aphasia. *Aphasiology*, *28*(1), 25-48.

MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. *Perception & Psychophysics*, *24*(3), 253-257.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in cognitive sciences*, *8*(5), 204-210.

Malfait, N., Fonlupt, P., Centelles, L., Nazarian, B., Brown, L. E., & Caclin, A. (2014). Different neural networks are involved in audiovisual speech perception depending on the context. *Journal of Cognitive Neuroscience*, *26*(7), 1572-1586.

Manuel, S. Y., Repp, B. H., Studdert-Kennedy, M., & Liberman, A. M. (1983). Exploring the "McGurk effect". *The Journal of the Acoustical Society of America*, *74*(S1), S66-S66.

Martin, A. D. (1974). Some objections to the term apraxia of speech. *Journal of Speech and Hearing Disorders*, *39*(1), 53-64.

Massaro, D. W., & Cohen, M. M. (1993). Perceiving asynchronous bimodal speech in consonant-vowel and vowel syllables. *Speech Communication*, *13*(1), 127-134.

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle* (Vol. 1). MIT Press.

Matchin, W., Groulx, K., & Hickok, G. (2014). Audiovisual speech integration does not rely on the motor system: Evidence from articulatory suppression, the Mcgurk effect, and fMRI. *Journal of Cognitive Neuroscience*, *26*(3), 606-620.

McGettigan, C., Faulkner, A., Altarelli, I., Obleser, J., Baverstock, H., & Scott, S. K. (2012). Speech comprehension aided by multiple modalities: Behavioural and neural interactions. *Neuropsychologia*, *50*(5), 762-776.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264* (5588), 746–748.

McNeil, M.R., Pratt, S.R., Fossett, T.R.D. (2004). The differential diagnosis of apraxia of speech. In: B, Maassen, editor. *Speech motor control in normal and disordered speech*. New York: Oxford University Press, 389–412.

Meredith, M.A., Nemitz, J.W., Stein, B.E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience,* 7, 3215–3229.

Meredith, M.A., Stein, B.E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research, 365*, 350–354.

Meredith, M.A., Stein, B.E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology, 75*, 1843–1857.

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

Mozaffarian D, Benjamin EJ, Go AS, et al. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. *Circulation*. 2015 ;e29-322.

Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., & Laurienti, P. J. (2012). Multisensory integration and aging.

Musacchia, G., Sams, M., Nicol, T., & Kraus, N. (2006). Seeing speech affects acoustic information processing in the human brainstem. *Experimental Brain Research*, 168, 1-10.

Musiek, F. E., Baran, J. A., & Pinheiro, M. L. (1992). P300 results in patients with lesions of the auditory areas of the cerebrum. *Journal of the American Academy of Audiology*, *3*(1), 5-15.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, *42*(4), 313-329.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1980). Brain potential correlates of voluntary and involuntary attention. *Progress in brain research*, *54*, 343-348.

Näätänen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological psychology*, *14*(1-2), 53-98.

Näätänen, R., & Alho, K. (1995). Mismatch negativity-a unique measure of sensory processing in audition. *International Journal of Neuroscience*, *80*(1-4), 317-337.

Näätänen, R., Lehtokoski, A., Lennest, M., Luuki, A., Alliki, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses.

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological bulletin*, *125*(6), 826.

Näätänen R. (2001); The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent MMNm. *Psychophysiology*, 38,1–21.

Näätänen, R, Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *118*(12), 2544–90.

Näätänen, R., Kujala, T., Escera, C., Baldeweg, T., Kreegipuu, K., Carlson, S., & Ponton, C. (2012). The mismatch negativity (MMN)–a unique window to disturbed central auditory processing in ageing and different clinical conditions.*Clinical Neurophysiology*, *123*(3), 424-458.

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage*, *59*(1), 781-787.

National Institute on Deafness and Other Communication Disorders (2015), retrieved from http://www.nidcd.nih.gov/health/voice/pages/apraxia.aspx

National Institute of Neurological Disorders and Stroke. (n.d.).NINDS aphasia information page. Retrieved from http://www.ninds.nih.gov/disorders/aphasia/aphasia.htm

New, A. B., Robin, D. A., Parkinson, A. L., Duffy, J. R., McNeil, M. R., Piguet, O., ... & Ballard, K. J. (2015). Altered resting-state network connectivity in stroke patients with and without apraxia of speech. *NeuroImage: Clinical*.

Nolfe, G., Cobianchi, A., Mossuto-Agatiello, L., & Giaquinto, S. (2006). The role of P300 in the recovery of post-stroke global aphasia. *European journal of neurology*, *13*(4), 377-384.

Ogar, J., Slama, H., Dronkers, N., Amici, S., & Luisa Gorno-Tempini, M. (2005). Apraxia of speech: an overview. *Neurocase*, *11*(6), 427-432.

Ohala, J.J. (1990).There is no interface between phonology and phonetics: A personal view. *Journal of Phonology, 18,*153–171.

Ojanen, V., Möttönen, R., Pekkola, J., Jääskeläinen, I. P., Joensuu, R., Autti, T., & Sams, M. (2005). Processing of audiovisual speech in Broca's area. *Neuroimage*, *25*(2), 333-338.

Park, W., Kwon, G. H., Kim, Y. H., Lee, J. H., & Kim, L. (2016). EEG response varies with lesion location in patients with chronic stroke. *Journal of neuroengineering and rehabilitation*, 13(1), 1.

Pettigrew, C., Murdoch, B., Kei, J., Ponton, C., Alku, P., & Chenery, H. (2005). The mismatch negativity (MMN) response to complex tones and spoken words in individuals with aphasia. *Aphasiology*, *19*(2), 131-163.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329-347.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., ... & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria.*Psychophysiology*, *37*(02), 127-152.

Pizzagali, D. A. **(**2007). Electroencephalography and high-density electrophysiological source localization. In J. T. Cacioppo,L. G. Tassinary, & G. G. Bernsten (Eds.), *The Handbook of Psychophysiology* (pp. 56–84). Cambridge, United Kingdom: Cambridge University Press.

Polich, J. (2003). Theoretical overview of P3a and P3b. in *Detection of Change*. pp 83-98

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–48.

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, *41*(2), 103–146.

Reisberg, D., Mclean, J., Goldfield, A., (1987). Easy to Hear but Hard to Understand: A Lip-Reading Advantage with Intact Auditory Stimuli.

Rosenfield, D.B. (1990). Speech apraxia in cortico-basal-ganglionic degeneration. *Annals of Neurology*,30: 296.

Rugg, M. D., & Coles, M. G. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford University Press.

Saint-Amour, D., De Sanctis, P., Molholm, S., Ritter, W., & Foxe, J. J. (2007). Seeing

voices: High-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia*, *45*(3), 587-597.

Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T., & Simola, J. (1991). Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience letters*, *127*(1), 141-145.

Sato, M., Buccino, G., Gentilucci, M., & Cattaneo, L. (2010). On the tip of the tongue: modulation of the primary motor cortex during audiovisual speech perception. *Speech Communication*, *52*(6), 533-541.

Schmid, G., & Ziegler, W. (2006). Audio–visual matching of speech and non-speech oral gestures in patients with aphasia and apraxia of speech. *Neuropsychologia*, *44*(4), 546-555.

Seddoh, S. A., Robin, D. A., Sim, H. S., Hageman, C., Moon, J. B., & Folkins, J. W. (1996). Speech timing in apraxia of speech versus conduction aphasia.*Journal of Speech, Language, and Hearing Research*, *39*(3), 590-603.

Sekuler, R., Sekuler, A.B., & Lau, R. (1997) Sounds alter visual motion perception. *Nature, 385,* 308.

Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *The Journal of the Acoustical Society of America*, *106*(2), 1078-1083.

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage*, *25*(1), 76-89.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, *17*(10), 2387-2399.

Solé, M.-J., Beddor, P. S., & Ohala, M. (2007). *Experimental Approaches to Phonology***.** Oxford, England: Oxford University Press.

Square-Storer, P., Darley, F. L., & Sommers, R. K. (1988). Nonspeech and speech processing skills in patients with aphasia and apraxia of speech. *Brain and Language*, *33*(1), 65-85.

Staiger, A., Finger-Berg, W., Aichert, I., & Ziegler, W. (2012). Error variability in apraxia of speech: A matter of controversy. *Journal of Speech, Language, and Hearing Research*, *55*(5), S1544-S1561.

Stein, B.E., Meredith, M.A., Huneycutt, W.S., & McDade, L. (1989). Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience, 1,* 12-24.

Stekelenburg, J. J., & Vroomen, J. (2012). Electrophysiological evidence for a multisensory speech-specific mode of perception. *Neuropsychologia*, *50*(7), 1425-1431.

Strand, E. A., & McNeil, M. R. (1996). Effects of length and linguistic complexity on temporal acoustic measures in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *39*(5), 1018-1033.

Stevenson, R. A., & James, T. W. (2009). Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage*, *44*(3), 1210-1223.

Strauss Hough, M., Downs, C. R., Cranford, J., & Givens, G. (2003). Measures of auditory processing in aphasia: Behavioural and electrophysiological analysis. *Aphasiology*, *17*(2), 159-172.

Sumby, W.H. & Pollack, I. (1954). The visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26,* 212.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187-1188.

Tervaniemi, M., Lehtokoski, A., Sinkkonen, J., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Test–retest reliability of mismatch negativity for duration, frequency and intensity changes. *Clinical Neurophysiology*, *110*(8), 1388-1393.

Tian, X., & Poeppel, D. (2012). Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation. *Mental Imagery*, 152.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952-981.

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological Science*, *21*(10), 1532-1540.

Tye-Murray, N., Sommers, M., Spehar, B., et al. (2008). Auditory-visual discourse comprehension by older and young adults in favorable and unfavorable conditions. *International Journal of Audiology, 47*, S31–S37.

Tye-Murray, M., Sommers, M., Spehar, B., Myerson, J., & Hale, S. (2010). Aging,

audiovisual integration, and the principle of inverse effectiveness. *Ear and Hearing, 31,* 636-644.

Uhlhaas, P. J., & Singer, W. (2006). Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron*, *52*(1), 155-168.

van der Merwe A. (1997); A theoretical framework for the characterization of pathological speech sensorimotor control. In: MR, McNeil, editor. *Clinical Management of Sensorimotor Speech Disorders*. New York: Thieme, 1–25.

Venezia, J.H., Fillmore, P., Matchin, W., Isenberg, A.L., Hickock, G., & Fridriksson, J. (2016). Perception drives production across sensory modalities: A network for sensorimotor integration of visual speech. *NeuoroImage, (126),* 196-207.

Wambaugh J.L., Doyle P. (1994). Treatment for acquired apraxia of speech: A review of efficacy reports. *Clinical Aphasiology*, 22: 231–243.

Wambaugh, J.L., Kaylinyak-Fliszar, M.M., West, J.E., Doyle, P.J.(1998). Effects of treatment for sound errors in apraxia of speech. *Journal of Speech Language Hearing Research*; 41: 725–743.

Wambaugh J. (2002). A Summary of treatments for apraxia of speech and review of replicated approaches. In: MR, McNeil, editor. *Seminars in speech and language. Apraxia of speech: From concept to clinic*. New York: Thieme, 293–308.

Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. A. (2006). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence.

Ward, D. & Scaler Scott, K. (2011). *Cluttering: A Handbook of Research, Intervention, and Education*. Psychology Press.

Watkins, K., & Paus, T. (2006). Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of cognitive neuroscience*, *16*(6), 978–87.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, *46*(3), 233–51.

Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., Aaltonen, O., et al. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*(3), 357–369.

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), 701–2.

Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: concepts and controversies. *Journal of Speech, Language, and Hearing Research*, *55*(5), S1485-S1501.

Appendix

TEACHERS COLLEGE
COLUMBIA UNIVERSITY

Teachers College IRB — *Continuing Review Approval Notification*

To: Melissa Randazzo
From: Curt Naser, TC IRB Administrator
Subject: IRB Approval: 15-297 Protocol
Date: 04/13/2016

Please be informed that as of the date of this letter, the Institutional Review Board for the Protection of Human Subjects at Teachers College, Columbia University has approved your *continuing* study, entitled "*Audiovisual Integration in Apraxia of Speech* on 04/13/2016."

The approval is effective until **04/12/2017**.

The IRB Committee must be contacted if there are any changes to the protocol during this period. **Please note:** If you are planning to continue your study, a Continuing Review report must be submitted to either close the protocol or request permission to continue for another year. Please submit your report by **03/15/2017** so that the IRB has time to review and approve your report if you wish to continue your study. The IRB number assigned to your protocol is **15-297**. Feel free to contact the IRB Office (212-678-4105 or IRB@tc.edu) if you have any questions.

Please note that your Consent form bears an official IRB authorization stamp. Copies of this form with the IRB stamp must be used for your research work. Further, all research recruitment materials must include the study's IRB-approved protocol number. You can retrieve a PDF copy of this approval letter from the Mentor site.

Best wishes for your research work.

Sincerely,
Curt Naser, Ph.D.
TC IRB Administrator

Attachments:
- b.Consent_cleancopy_22July2015.pdf