



University of Tennessee Health Science Center
UTHSC Digital Commons

Theses and Dissertations (ETD)

College of Graduate Health Sciences

5-2017

Automatic Activation of Phonological Templates for Native but Not Nonnative Phonemes: An Investigation of the Temporal Dynamics of Mu Activation

Daniela Cristina Santos-Oliveira
University of Tennessee Health Science Center

Follow this and additional works at: <https://dc.uthsc.edu/dissertations>

 Part of the [Speech and Hearing Science Commons](#)

Recommended Citation

Santos-Oliveira, Daniela Cristina (<http://orcid.org/0000-0002-1061-0801>), "Automatic Activation of Phonological Templates for Native but Not Nonnative Phonemes: An Investigation of the Temporal Dynamics of Mu Activation" (2017). *Theses and Dissertations (ETD)*. Paper 437. <http://dx.doi.org/10.21007/etd.cghs.2017.0431>.

This Dissertation is brought to you for free and open access by the College of Graduate Health Sciences at UTHSC Digital Commons. It has been accepted for inclusion in Theses and Dissertations (ETD) by an authorized administrator of UTHSC Digital Commons. For more information, please contact jwelch30@uthsc.edu.

Automatic Activation of Phonological Templates for Native but Not Nonnative Phonemes: An Investigation of the Temporal Dynamics of Mu Activation

Document Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Program

Speech and Hearing Science

Track

Speech-Language Science

Research Advisor

Devin Casenhiser, Ph.D.

Committee

Ferenc Bunta, Ph.D. Ashley W Harkrider, Ph.D. Mark S Hedrick, Ph.D. Tim Saltuklaroglu, Ph.D. Antonio J. S. Teixeira, Ph.D.

ORCID

<http://orcid.org/0000-0002-1061-0801>

DOI

10.21007/etd.cghs.2017.0431

**Automatic Activation of Phonological Templates for Native but Not Nonnative
Phonemes: An Investigation of the Temporal Dynamics of Mu Activation**

A Dissertation
Presented for
The Graduate Studies Council
The University of Tennessee
Health Science Center

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
From The University of Tennessee

By
Daniela Cristina Santos-Oliveira
May 2017

Copyright © 2017 by Daniela Cristina Santos-Oliveira.
All rights reserved.

DEDICATION

This dissertation is dedicated to my family for their unconditional support throughout this long journey, specially to my two daughters for being my light and motivation.

ACKNOWLEDGEMENTS

I would like to thank Dr. Devin Casenhiser for his mentorship and guidance. I appreciate the time and effort he has invested in studying research topics not directly in line with his interests to help me pursue my own research work and ultimately this dissertation. I am grateful for his support.

I would like to thank all my committee members, Dr. Ferenc Bunta, Dr. Ashley Harkrider, Dr. Mark Hedrick, Dr. Tim Saltuklaroglu, and Dr. Antonio Teixeira for their patience, advice, criticism and understanding throughout the long road leading to this dissertation work. A special note of gratitude to Dr. Ashley Harkrider for allowing me to pursue my dissertation work in her lab.

I would like to thank my lab colleagues, David Thornton, David Jenson and Tiffani Kittilstved, for the time they expended on me helping to achieve this goal. I am especially grateful to David Thornton for his guidance in learning all the lab routines, insightful discussions, and comforting presence, making this dissertation work possible.

I would like to thank the Portuguese national funding agency for science, research and technology (www.fct.pt) for granting me a scholarship to support the completion of the Ph.D. program.

I would also like to thank all my friends and family for their encouragement, lending me a comforting ear when I was down and laughing with me in times of joy. Thank you for wise counsel, support, and comic relief in times of stress.

Finally, a very special note of gratitude to my parents and my husband. I would not have make it here without your support and patience.

ABSTRACT

Models of speech perception suggest a dorsal stream connecting the temporal and inferior parietal lobe with the inferior frontal gyrus. This stream is thought to involve an auditory-motor loop that translates acoustic information into motor/articulatory commands and is further influenced by decision making processes that involve maintenance of working memory or attention. Parsing out dorsal stream's speech specific mechanisms from memory related ones in speech perception poses a complex problem. Here I argue that these processes may be disentangled from the viewpoint of the temporal dynamics of sensorimotor neural activation around a speech perception related event.

Methods: Alpha (~10Hz) and beta (~20Hz) spectral components of the mu (μ) rhythm, localized to sensorimotor regions, have been shown to index somatosensory and motor activity, respectively. In the present work, event related spectral perturbations (ERSP) of the EEG μ -rhythm were analyzed, while manipulating two factors: active/passive listening, and perception of native/nonnative phonemes. Active and passive speech perception tasks were used as indexes of memory load employed, while native and nonnative perception were used as indexes of automatic top-down coding for sensory analysis.

Results: Statistically significant differences were found in the oscillatory patterns of μ components between active and passive speech perception conditions with greater μ alpha and beta event related desynchronization (ERD) after stimuli offset in active speech perception. When compared to listening to noise, passive speech perception presented significantly ($pFDR < 0.05$) stronger alpha and beta ERD during and after stimuli presentation. When comparing native to nonnative speech perception, stronger alpha (8-14Hz) and beta (22-25Hz) event related synchronization (ERS) were observed before and during stimuli onset in the passive nonnative task. Passive native perception, on the other hand, presented stronger alpha and especially beta ERD before stimuli onset as well as stronger alpha ERD between presentation of the two syllables composing the stimuli (600-850ms), and during the presentation of the second syllable (1000ms).

Conclusion: These findings suggest that neural processes within the dorsal auditory stream are functionally and automatically involved in speech perception mechanisms. While its early activity (shortly after stimuli onset) seems to be importantly involved with the instantiation of predictive motor/articulatory internal models that help constraining speech discrimination, its later activity (post-stimulus offset) seems essential in the maintenance of working memory processes.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION	1
CHAPTER 2. LITERATURE REVIEW	7
Active and Passive Speech Perception	7
Perception of Nonnative Categories	10
Models of Sensorimotor Integration in Speech Processing	13
EEG and Sensorimotor Integration	17
Independent Component Analysis (ICA).....	18
The Arceu or μ -Rhythm.....	19
Event-Related Synchronization and Desynchronization.....	19
CHAPTER 3. METHODS AND PROCEDURES	22
Methods	22
Materials	22
Participants.....	22
Procedures.....	24
Data Acquisition	25
Data Processing.....	25
Individual EEG data processing/analyses.....	27
Group EEG data processing/analyses	28
CHAPTER 4. RESULTS.....	30
Discrimination Accuracy	30
μ Clusters	30
Event-Related Spectral Perturbations	30
CHAPTER 5. DISCUSSION	41
Experimental Results	41
Limitations and Future Directions	44
LIST OF REFERENCES.....	46
VITA.....	61

LIST OF FIGURES

Figure 3-1. Trial timeline.....	23
Figure 3-2. Stimuli presentation timeline.	26
Figure 4-1. Cluster results for the left-hemisphere μ components.	31
Figure 4-2. Cluster results for the right-hemisphere μ components.	32
Figure 4-3. Mean left and right hemisphere μ time-frequency ERSP's.....	33
Figure 4-4. Mean left and right hemisphere ERSP's for passive (native and nonnative) versus active (native and nonnative) perception tasks.	35
Figure 4-5. Mean left and right hemisphere ERSP's for passive native versus active native perception tasks.	36
Figure 4-6. Mean left and right hemisphere ERSP's for passive nonnative versus active nonnative perception tasks.....	37
Figure 4-7. Mean left hemisphere ERSP's for passive native versus passive nonnative speech perception tasks.	38
Figure 4-8. Mean left and right hemisphere ERSP's for listening to noise versus passive native speech perception task.	39
Figure 4-9. Mean left and right hemisphere ERSP's for listening to noise versus passive nonnative speech perception task.	40

CHAPTER 1. INTRODUCTION

Speech perception is a dynamic and essentially multisensory neural process, which has been shown to involve several areas of the brain. These areas include temporal regions, such as the superior temporal gyrus (STG) and the superior temporal sulcus (STS), which are traditionally implicated in auditory analysis (Belin, Zatorre, & Ahad, 2002; Okada et al., 2010), and also frontal regions, such as the inferior frontal gyrus (IFG) and the premotor cortex (PMC) that are usually associated with converting auditory percepts to articulatory representations of speech (Hickok & Poeppel, 2015). This concept of speech perception as a highly dynamic auditory-motor system dates back to Wernicke's conception of the neural network behind language processing (Wernicke, 1969).

Building on Wernicke's concept, Hickok and Poeppel (2004, 2007) proposed a comprehensive model of speech perception, the dual stream model, that accounts for the dynamic interaction between auditory and sensorimotor processes. According to this model, speech perception starts by involving auditory-responsive fields in STG bilaterally, then diverting into two processing streams: a ventral stream, and a dorsal stream. The ventral stream originates in the upper posterior part of the temporal lobe and extends toward the anterior part of the temporal lobe, where it also connects to the ventral inferior frontal gyrus through the uncinate fasciculus and extreme capsule, and it has been shown to be involved in mapping sound onto meaning (Hickok & Poeppel, 2007; Rauschecker & Scott, 2009; Saur et al., 2008; Scott & Wise, 2004; Specht, 2013; Weiller, Bormann, Saur, Musso, & Rijntjes, 2011). The dorsal stream extends from the posterior temporal lobe through inferior parietal areas into the left inferior frontal gyrus, including premotor areas¹. This stream mainly follows the arcuate and longitudinal fasciculi, connecting the temporal and inferior parietal lobe with the inferior frontal gyrus (Catani et al., 2007; Houde & Nagarajan, 2011). Traditionally, the dorsal stream has been reported to be left lateralized. More recently, however, evidence has been found to suggest a more bilateral organization in speech (Cogan et al., 2014; Simmonds et al., 2014). Hickok and Poeppel suggest that the dorsal stream is an auditory-motor loop that translates acoustic information into motor/articulatory commands (Hickok & Poeppel, 2004, 2007; Rauschecker & Scott, 2009; Scott & Wise, 2004). Hickok et al. (2011) have, in fact, proposed a framework of speech sensorimotor integration which helps explain the activation and possible top-down modulatory influence of the motor system during speech perception. They view it as a feedback control architecture by which inverse internal models convert the input auditory signal into potential articulatory planning control signals; the auditory consequences of these articulatory planning control signals are then predicted by forward internal models that are compared with the auditory speech representation. These internal models are constructed based on previous experience or learned associations between the motor representations of a phonological signal and their sensory consequences (Hickok et al., 2011).

¹ Rauschecker and Scott (2009) have specified the superior and dorsal premotor cortex as the premotor regions being influenced by the dorsal stream.

Inconsistent research findings regarding dorsal stream activity, however, have muddied the interpretation of findings in relation to its function in speech perception. (Chevillet, Jiang, Rauschecker, & Riesenhuber, 2013; Leech, Holt, Devlin, & Dick, 2009; E. B. Myers, Blumstein, Walsh, & Eliassen, 2009). Consistent with the Hickok and Poeppel's model, a number of studies suggest that activation within the dorsal stream reflects the automatic comparison of sensory information from speech against an internal phonological or articulatory template (automatic predictive coding or hypothesis testing) (D. Callan, Callan, Gamez, Sato, & Kawato, 2010; M. Iacoboni, 2008; Kawato, 1999; Stephen M. Wilson & Iacoboni, 2006). However, because the dorsal stream has also been found to be active during decision making processes that involve maintenance of working memory or attention (Binder, Liebenthal, Possing, Medler, & Ward, 2004; Buchsbaum & D'Esposito, 2008; Krawczyk, 2002), it has been difficult to separate out primarily speech-related neural activation from activation that is chiefly related to attention and working memory when examining the relative magnitude of neural activation associated with these processes. These processes may, however, be distinguishable from the viewpoint of the dynamics of activation across the time-course of speech processing.

In particular, a number of studies have demonstrated that activation of the dorsal stream varies with the cognitive demands of the task. Conditions where subjects are actively asked to discriminate speech sounds involve the maintenance of those sounds in working memory to facilitate the comparison of those sounds. On the other hand, passive conditions, in which subjects passively listen to speech do not require the maintenance of sounds in working memory and therefore are less cognitively demanding. Thus, research has shown that tasks requiring active discrimination of phonemes show greater dorsal stream activation than passive listening to the same phonemes (Jussi Alho et al., 2014; Daniel Callan, Callan, & Jones, 2014; Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007). Similarly, discrimination of phonemes in noisy conditions (while intelligible) evidences a more significant involvement of motor regions, within the dorsal stream, in comparison to discrimination of phonemes in quiet conditions (Binder et al., 2004; D'Ausilio, Bufalari, Salmas, & Fadiga, 2012; Du, Buchsbaum, Grady, & Alain, 2014; Jenson et al., 2014). Thus the cognitive demands of the task which require greater effort and/or greater reliance on a phonological working memory seem to elicit greater activity in the dorsal stream.

These findings, however, do not obviate the notion that the dorsal stream is involved in the automatic activation of articulatory templates if one considers the dynamic nature of the time-course of activation. Dorsal stream activation related to task reliance on a phonological working memory buffer should be observed to happen after stimuli have been presented. Dorsal stream activity related to automatic activation of articulatory templates, on the other hand, should occur within 200ms of hearing the stimulus (D. Callan et al., 2010). Thus it is not the absolute or relative amount of dorsal stream activity alone, but how that activation changes across the time-course of processing that is important when distinguishing the functions of dorsal stream activity in active versus passive tasks.

Much of the research cited in support of one or the other role of the dorsal stream, however, is not designed to measure the time-course of activation; instead, it measures the average activation across entire trials. For example, Callan et al. (2010) examined the role of the PMC in perceptual performance under noise, using fMRI and MEG in both active discrimination and passive listening tasks. The authors found PMC activation across all conditions with greater activation during active tasks, and during correct trials. They interpreted the existence of dorsal stream activity during passive trials as indicating an automatic role for this area, and suggested the greater activity during active and correct trials reflected greater reliance on phonological working memory. While we tend to agree with this interpretation, assumedly both passive and active conditions embody some measure of attention. It could therefore be argued that the existence of lower levels of dorsal stream activation during passive listening indicates lower levels of attention being paid to the stimuli, and is not on its own an indication of activation of automatic sensorimotor processing. The time-course of activation would in this case help to parse out these effects. If the greater activation associated with active trials were to occur primarily after stimuli offset, this would suggest that the activation is indeed due to greater reliance on the phonological memory buffer.

A second way in which one might attempt to disentangle the effects of cognitive load on dorsal stream activity in speech perception is by comparing nonnative phoneme activation to native phoneme activation. Because nonnative phonemes are not part of the articulatory repertoire of a speaker, they should elicit less automatic dorsal stream activation during speech processing (in comparison to native phonemes), and potentially more activation after stimuli offset as an indication of the greater effort required to process the unfamiliar phonemes. An fMRI study by Callan et al. (2003) suggests that nonnative phonemes may well elicit less dorsal stream activity. Their results indicated significantly less dorsal stream activity for nonnative phonemes prior to discrimination training in comparison to dorsal stream activity after discrimination training. The authors suggest the post-training dorsal stream activity is due to the formation of auditory-articulatory (perceptual-motor) mappings related to speech processing and that formed as a result of the phoneme discrimination training that took place. Prior to training no such mappings existed resulting in markedly less automatic auditory-articulatory processing. Although Callan et al. (2003) did not investigate timing differences, and did not make a direct comparison between native and nonnative speech sounds, their results suggest that the time-course of dorsal stream activation should show greater sensorimotor activity close after stimuli onset during the perception of native sounds, and comparatively less for the perception of nonnative sounds.

One methodology well-suited to examining the combination of these factors in a single experiment is the analysis of spectral temporal activity derived from analysis of electroencephalogram (EEG) oscillations. Research has shown that measuring changes in spectral power within the mu (μ) rhythm, whose pattern is characterized by oscillatory peaks in both alpha (~ 10 Hz) and beta (~ 20 Hz) frequency bands, is a good method for understanding sensorimotor processes related to speech perception (A. Bowers, Saltuklaroglu, Harkrider, & Cuellar, 2013; A. L. Bowers, Saltuklaroglu, Harkrider, Wilson, & Toner, 2014; D. Callan et al., 2010; Jenson et al., 2014; Skipper, van

Wassenhove, Nusbaum, & Small, 2007). The μ alpha component frequencies of the sensorimotor μ -rhythm tend to originate from the somatosensory cortex when guidance is needed for ongoing movement, while its μ beta component frequencies tend to emerge in a somatotopic manner from the precentral gyrus corresponding with the motor cortex for the effector involved (i.e., lip vs. hand movements) and indicate motor activity (Hari, 2006; Jensen et al., 2005; Jaime A. Pineda, 2005). Therefore, the analysis of patterns within μ -alpha and μ -beta frequencies around speech related events is likely to disclose important information regarding the timing of sensorimotor integration processes in speech processing.

Measures of μ -rhythm oscillatory activity can be visualized using event-related spectral perturbations (ERSPs) which allow for analysis of changes in spectral power across a range of frequencies. Changes in spectral power are thought to indicate degrees of synchronization of the action potentials of neurons in the area being investigated. Decreases in spectral power reflect event-related desynchronization (ERD) of cortical neural network oscillations, while increases in spectral power reflect event-related synchronization (ERS) of cortical neural network (Pfurtscheller & Lopes da Silva, 1999).

ERD and ERS are band and location specific, a fact which has resulted in several interpretations of their cognitive correlates. Of particular interest to the present work is that alpha ERD is thought to indicate a release from inhibitory sensory filtering or gating and may also contribute to predictive coding (W. Klimesch, 2012). Beta ERD, on the other hand, is often associated with predictive ‘top-down’ coding (hypothesis testing) for sensory analysis (Buschman & Miller, 2007; Kilavik, Zaepffel, Brovelli, MacKay, & Riehle, 2013; Siegel, Donner, & Engel, 2012). Monotonic alpha power has also been found to decrease with increased attentional or cognitive demands. Weisz and colleagues (2011) study indicated that a monotonic decrease in alpha power (i.e., alpha ERD) across a wide network of brain regions (prefrontal, temporal, parietal) was related both to reductions in speech intelligibility and working memory set size, supporting its sensitivity to both speech perception and working memory. μ ERS, on the other hand, has been interpreted as having inhibitory or gating effects which can be very important to optimize energy demands and control over excitatory processes (W. Klimesch, 2012; W. Klimesch et al., 1996). Research has connected μ ERS with gating of noise in speech discrimination tasks (Jenson et al., 2014).

In addition, research has shown that μ -alpha ERD in speech production may provide an index of sensory feedback in audio-vocal monitoring (Jenson et al., 2014; Jenson, Harkrider, Thornton, Bowers, & Saltuklaroglu, 2015; Tamura et al., 2012). Recent studies analyzing oscillations of the sensorimotor μ rhythm in both speech perception and production (Jenson et al., 2014; Jenson et al., 2015) have contributed with further evidence that μ alpha and μ beta can be reliable indexes of sensory feedback and forward predictive models (hypothesis testing), respectively, in both speech perception and production.

Consequently, given that μ rhythm is localized to primary motor/PMC regions, and that the PMC is strategically localized within the auditory dorsal stream, maintaining

a bidirectional communication between auditory and somatosensory regions via the arcuate and longitudinal fasciculi, μ alpha and beta oscillations present a unique system for the analysis of the phonological sensory loop mechanisms (μ beta, predictive forward model and μ alpha, sensory feedback). In fact, a recent body of research analyzing the time course activity of the μ rhythm (A. Bowers et al., 2013; A. L. Bowers et al., 2014; Jenson et al., 2014; Jenson et al., 2015) strongly suggests that the dorsal auditory stream supports a variety of roles in speech perception and production processes (articulatory representations, attention, experience and memory load).

Utilizing high temporal resolution techniques, like EEG, to analyze neural activity behind speech perception has, therefore, been shown to be fruitful. In the past, problems associated with artifact removal and the biophysical inverse problem used to limit its reach and application. Recently, however, new approaches to EEG data analysis have been found to bypass these problems. ICA is an algorithm-based analysis that applied to EEG data aims to identify spatially fixed and temporally independent sources of neural activity that are linearly mixed across several sensors (A. Delorme & Makeig, 2004; Makeig, Debener, Onton, & Delorme, 2004; Onton, Westerfield, Townsend, & Makeig, 2006). It can, for example, separate out artifacts embedded in the data (e.g.: eye movements, eye blinks, muscle tension), since they are usually independent of each other. Applying ICA to EEG data brings together high time/frequency resolution with enhanced spatial resolution.

Consequently, the analysis of μ alpha and μ beta ERSPs from EEG activity via ICA represents an ideal methodology through which to explore both the nature and timing of sensorimotor processes in the perception of speech. More specifically, the suggested method is able to verify whether phonological activation related to working memory occurs both after stimulus presentation, and during active rather than passive tasks. Similarly, since nonnative phonemes lack entrenched articulatory templates, contrastive analysis of native and nonnative phonemes can verify whether dorsal stream activity occurring during stimulus presentation is related to automatic phonological processing. Moreover, since nonnative phonemes may well require more attention than native ones, crossing this factor with the active/passive conditions as we suggest above, allows us to disentangle the effect of working memory from the effect of automatic phonological processing.

In the present study we hypothesize that activation related to automatic phonological processing will occur during stimulus presentation while processing related to working memory will occur after stimulus presentation. To test these hypotheses, we investigate sensorimotor μ -rhythm activity localized to the dorsal stream during speech perception while manipulating two factors: native/nonnative phonemes and active/passive listening. This produces the following conditions: 1) listening to noise (control); 2) passively listening to a native speech contrast; 3) active discrimination of a native speech contrast; 4) passively listening to a nonnative speech contrast; and 5) active discrimination of a nonnative speech contrast. The hypotheses tested are the following:

- **H1:** Greater sensorimotor activity after stimuli offset is correlated with working

memory buffer processes, instantiated within the dorsal auditory stream. If this is the case, greater sensorimotor activity (μ -alpha and μ -beta ERD) after stimuli offset is expected in active discrimination tasks in comparison to both passive and noise conditions.

- **H2:** Sensorimotor activation during stimuli presentation is functionally related to the automatic instantiation of auditory-articulatory mappings (entrenched for native phonemes, but not for nonnative phonemes. If this is the case, we will find: (1) greater sensorimotor activity (μ -alpha and μ -beta ERD) during passive stimulus presentation for the native phonemes in comparison to nonnative phonemes, and (2) greater activation in both speech conditions (native and nonnative) in comparison to noise. Early (within 200ms of stimuli onset) sensorimotor activity during stimuli presentation in passive conditions and not later on suggests involvement in automatic speech related processes rather than processes related to working memory buffer.

CHAPTER 2. LITERATURE REVIEW

Activity within the dorsal auditory stream has extensively been argued to relate to the maintenance of working memory and decision-making processes (J. Alho et al., 2012; Binder et al., 2004; Buchsbaum & D'Esposito, 2008; Holt & Lotto, 2010; Krawczyk, 2002). There is however a great deal of research also involving this stream activity with the generation of auditory-articulatory mappings implicated in the predictive coding of the articulatory consequences of sensory information in speech perception (D. Callan et al., 2010; Hickok & Poeppel, 2007; M. Iacoboni, 2008; Kawato, 1999; Stephen M. Wilson & Iacoboni, 2006). These two processes are generally highly interconnected and their disentanglement can be hard to achieve. Therefore, most research exploring the role of the dorsal auditory stream in speech perception safeguards both possibilities. The current project ventures a complex research paradigm to try the disentanglement of these two functions of the dorsal auditory stream during speech perception. By the analysis of μ alpha and μ beta ERSPs from EEG activity via ICA the current project intends to explore the time-course activity of sensorimotor processes involved in speech perception, more precisely during the active and passive perception of native and nonnative sounds. This complex approach is new and therefore the following review attempts to justify why the time-course analysis of native and nonnative perception, during active and passive tasks will contribute to a better understanding of the sensorimotor integration processes behind speech perception.

Active and Passive Speech Perception

One possible, often reported, confound related to the observation of activation within the dorsal auditory stream during speech perception and categorization tasks is this stream's involvement in the maintenance of working memory, attentional resources, and decision processes (J. Alho et al., 2012; Binder et al., 2004; Holt & Lotto, 2010). There are doubts whether its activation is due to attentional/decision-making processes involved with task performance or whether it is functionally related to speech perception and categorization processes. For this reason, the functional role of the dorsal auditory stream in speech perception and categorization keeps being debated and explored. A number of researchers have suggested that activation within the dorsal auditory stream is secondary to speech perception (Binder et al., 2004; D. Callan et al., 2010; D. E. Callan, Jones, Callan, & Akahane-Yamada, 2004; D'Ausilio et al., 2012; Hickok & Poeppel, 2007) – related to mental rehearsal and decoding degraded or ambiguous signal or even related to perception of nonnative speech sounds – while others have argued for a linguistic role of those areas (Chevillet et al., 2013; Y. S. Lee, Turkeltaub, Granger, & Raizada, 2012; Liebenthal, Sabri, Beardsley, Mangalathu-Arumana, & Desai, 2013; Meister et al., 2007).

This debate has also been ignited by conflicting results reported in studies using transcranial magnetic stimulation (TMS). In one hand, there are studies where stimulation of motor areas has shown to influence speech perception abilities, like categorical perception (R. Mötönen & Watkins, 2009), discrimination of specific places of

articulation (R. Möttönen & Watkins, 2012), or discriminating stop-consonants (Meister et al., 2007). Moreover, another, more recent study from Möttönen et al. (2013) has combined TMS and EEG to measure the effect of stimulation of the lip and hand representations of the left motor cortex on the MMN response. Their results revealed a reduced MMN response after TMS of the lips, but not of the hand representation. As noted by Specht (2014), this is a strong indication that disturbances within the articulatory network can automatically perturb speech sound discrimination. On the other hand and concurrently, there are studies using TMS and lesion data reporting that disturbances to the motor cortex have only minor or even no effects on general speech perception and comprehension abilities (Hickok, 2009, 2012; Lotto, Hickok, & Holt, 2009; Rogalsky, Love, Driscoll, Anderson, & Hickok, 2011).

Neurophysiological data derived from imaging studies has not fundamentally helped resolving this question either, mainly because of divergence in the research methods employed. A group of studies have, for example, employed research methods where only active perceptual discrimination and categorization tasks were used, i.e., tasks where the participants are explicitly asked to discriminate between speech sounds (Daniel E. Callan, Jones, & Callan, 2014; D. E. Callan et al., 2004; Du et al., 2014; Liu & Jin, 2015; Pulvermüller et al., 2006). As an example, Callan et al. (2014) have explored the differential involvement of speech motor brain regions during the perception of native and foreign-accented phonemes for first and second language listeners. They presented their participants, English native speakers and Japanese native speakers, with syllables starting either with /r/ or /l/, in both an unaccented condition (/r/ and /l/ spoken by native English speakers), and a foreign-accented condition (/r/ and /l/ spoken by native Japanese speakers). Participants had to actively identify the syllables as starting with either /r/ or /l/. In what respects specifically to perception of the foreign-accented condition, which was similarly difficult for both English and Japanese speakers (as judged by behavioral testing), results showed significant activation of the right inferior ventral PMC/Broca's area, the superior ventral and dorsal PMC, and the left cerebellum. Authors ruled out the hypothesis of PMC activation being related with general cognitive processes, since they controlled for task difficulty. However, another way this can be controlled for is by the use and contrast of active and passive speech perception tasks.

In relation to this idea, there is another body of research that seems to support the automatic and functional role of the dorsal auditory stream in speech perception and categorization that typically uses methods where the participants are distracted from the speech perception task at hand, by being asked to pay attention to other properties of the stimuli (Chevillet et al., 2013; Y. S. Lee et al., 2012), or use a controlled combination of passive and active tasks (Jussi Alho et al., 2014; D. Callan et al., 2014; Meister et al., 2007). Passive tasks are those in which participants are passively listening to the speech stimuli, therefore not being asked to make any decisions about the speech sounds heard. In fact, Wilson et al. (2004) have shown evidence that passively listening to speech may, in fact, activate motor areas involved in speech production (see also Benson et al., 2001; Binder et al., 2000). Using fMRI, they compared neural activation during the following tasks: 1) passively listening to meaningless monosyllables, 2) speech production (production of the same meaningless monosyllables), 3) listening to white noise or bell

sound (control task), and 4) bimanual motor task (control task). Their observations highlighted the activation of areas of the PMC for both speech perception and production, with diminished response to non-speech stimuli. The superior ventral PMC showed an enhanced response to speech stimuli, while the inferior ventral PMC was stronger for speech production. The superior ventral PMC showed a diminished response to non-speech sounds. As noted by the authors, these findings are “consistent with the view that speech perception involves the motor system in a process of auditory-to-articulatory mapping to access a phonetic code with motor properties” (S. M. Wilson et al., 2004, p. 702). In agreement with these findings, in a later study, Wilson and Iacoboni (2006) analyzed the involvement of premotor brain areas in relation to nonnative phonemes varying in producibility, and once more found PMC activation, bilaterally, while participants were passively listening to the speech stimuli. This study did not include an active discrimination task, but the activity was found in comparison to rest.

In order to contribute to the disentanglement of the involvement of the dorsal auditory stream in speech perception, as possibly having an essential role or instead acting as an aiding agent, the current project employs a research method where both active perceptual discrimination and passive listening to speech tasks are used. Besides, a control task in which participants only listen to white noise is also employed. Less activity within the dorsal auditory stream is expected in the passive tasks, when compared to active discrimination of native and nonnative phonemes, since the participants are not being asked to produce any judgments about the stimuli and, therefore, no modeling or motor planning of the sounds is expected to occur. Furthermore, in accordance with ‘dual-stream model’s’ assumptions, passive speech perception should elicit activity in the motor system only after stimulus onset, since forward models originated in PMC are thought to modulate perception by predicting the likely sensory consequences (Hickok & Poeppel, 2004, 2007; Skipper et al., 2007). Additionally, if dorsal auditory stream areas are automatically involved in speech perception, as opposed to non-speech signals, we should observe differences between passive speech perception tasks and listening to noise. PMC activity is expected in passively listening to the speech stimuli after stimulus onset, but not when listening to noise, in accordance with Wilson et al.’s results (2006; 2004) and the view that speech perception involves the motor system in a process of mapping auditory signal to the articulatory properties of the speech sounds.

In active perception, i.e. active discrimination of speech categories, ‘dual-stream model’ assumptions suggest that forward predictions generated in the PMC may generate predictions or articulatory hypotheses that will reduce the load on the sensory system by limiting the number of candidate perceptual targets. This model, in consequence, predicts early activity in the motor system and simultaneous activity in the sensory system for generation of those predictive articulatory hypothesis or expectations in regards to the sensory outcome, limiting the load for sensory analysis. This will result in activity within the dorsal auditory stream prior to speech stimuli onset, and possibly after stimulus offset, reflecting an initial articulatory hypothesis followed by synthesis with sensory information (i.e., sensorimotor integration).

Perception of Nonnative Categories

Activation of brain areas within the dorsal auditory stream associated with speech perception and phonological categorization tasks has been reported somewhat extensively in the literature (D. Callan et al., 2010; D. Callan et al., 2014; D. E. Callan et al., 2004; D. E. Callan, Tajima, et al., 2003; Hickok et al., 2011; Hickok & Poeppel, 2015; M. Iacoboni, 2008; Stephen M. Wilson & Iacoboni, 2006). The function of these areas in speech perception and phonological categorization, however, is still not fully understood. Different researchers have suggested that brain motor areas, including the PMC, could be involved in transforming the acoustic signal to a phonetic code or, more specifically, in generating online internal forward models of auditory templates or native phonemes (D. Callan et al., 2010; D. E. Callan et al., 2004; D. E. Callan, Jones, et al., 2003; M. Iacoboni, 2008; Poeppel, Idsardi, & van Wassenhove, 2008; Skipper et al., 2007; Stephen M. Wilson & Iacoboni, 2006). These internal forward models are formulated as being mechanisms that predict the sensory consequences of internally stimulated articulatory models in way to facilitate speech perception (see Models of Sensorimotor Integration in Speech Processing). It has been argued that the dorsal auditory stream, believed to instantiate these predictive mechanisms, is especially important in aiding speech perception when the speech signal is degraded either by noise or in nonnative speech perception (mainly when learning new phonological contrasts) (A. M. Callan, Callan, Tajima, & Akahane-Yamada, 2006; D. Callan et al., 2014; D. E. Callan et al., 2004; D. E. Callan, Tajima, et al., 2003; Pillai et al., 2003; Wang, Sereno, Jongman, & Hirsch, 2003; Zhang et al., 2009). Wilson and Iacoboni (2006), for example, used fMRI to analyze neural responses to nonnative phonemes varying in producibility and found activation in both superior temporal (auditory) and PMC areas to distinguish native and nonnative phonemes, with greater signal changes for nonnative phonemes. Furthermore, in line with Callan's (D. E. Callan et al., 2004) proposed model of forward internal predictive mechanisms, they suggest that internal representations of known phonemes result from the integration of sensory and motor systems, where the PMC is involved in generating internal forward models, while the auditory system is essentially responsible for "comparing the acoustic input to predicted acoustic consequences of phonemes under consideration" (Stephen M. Wilson & Iacoboni, 2006, p. 322). Their findings regarding the involvement of superior temporal areas in matching auditory input to stored templates (Hickok & Poeppel, 2004; Scott & Wise, 2004; Warren, Wise, & Warren, 2005) as well as their findings regarding the motor system role in the online generation of these internal templates (Hickok et al., 2011) find even further support in the literature and seem to be coherent with more constructivist theories of speech perception, which predict greater PMC engagement in the predictive coding of more difficult tasks. According to Wilson and Iacoboni's (2006) interpretation of their own results, the greater PMC activation associated with the presentation of nonnative phonemes is related to this area's engagement in repeated attempts to model the unfamiliar phonemes, when a match could not be obtained (templates can only be retrieved for native phonemes).

Further information regarding the time-course of activity of these sensorimotor areas in the discrimination of native and nonnative phonemes is extremely important to a full understanding of the involvement and function of the dorsal auditory stream in

speech perception. The present study will try to disentangle working memory load and decision making processes from more automatic phonological processes associated with the dorsal auditory stream by analyzing the time-course of sensorimotor activity within this stream in relation to the timing of stimuli presentation – native and nonnative phonological contrasts.

In the past, neural processing differences for native and nonnative phonemes have also been reported in several neurophysiological studies using mismatch negativity (MMN)². These studies, however, tend to focus on the function of the auditory cortex in speech perception and not so much on the involvement of motor regions in speech recognition. Nevertheless, MMN studies have contributed significantly to support evidence on the existence of a language-specific memory trace associated with phonemic coding and storage. As an example of two essential studies that led to this conclusion, Näätänen and colleagues, in 1997, found that in a sequence of native phonemes, deviant native phonemes (i.e., a native phoneme different from the one previously being presented) produced a larger MMN in the left hemisphere, localized in the superior temporal regions, than deviant nonnative phonemes (Näätänen et al., 1997). In the same year, Dehaene-Lambertz found a larger MMN produced by acoustic changes that resulted in a crossing between phonemic categories than equivalent acoustic changes within a same phoneme category, localized in the same auditory temporal regions of the left hemisphere (Dehaene-Lambertz, 1997). These results lead to the theoretical assumption of the existence of a memory trace specific to language, where relevant linguistic traits of the native phonological system were stored, assumedly in the superior temporal brain regions.

Additional evidence to the hypothetical existence of a language-specific memory trace associated with phonemic coding and storage came from training studies, where increased MMN were observed in the left hemisphere for newly learned phonetic categories (see Tremblay, Kraus, Carrell, & McGee, 1997). The greater MMN for native phonemes or learned contrasts consistently reported by MMN studies seems to contradict more recent findings using neuro imaging techniques, which report greater activity for nonnative phonemes. An explanation for the different results may however be related to differences in the stimuli used. The stimuli in Näätänen's MMN studies often consists of native phonemes that are discriminable as deviants, whereas the nonnative phonemes are frequently not discriminable from the standards. Conversely, in the neuroimaging studies nonnative stimuli are readily perceivable as nonnative or non-prototypical sounds. Importantly, despite possible methodological incongruences, the hypothesis of a language specific memory trace for phonemic categories raised by the late MMN studies seems to link well with the more recent literature indicating that phonological categorization of native sounds seems to rely more heavily in the superior temporal regions, consistent with increased reliance on auditory-phonetic representations (D. E. Callan et al., 2004),

² As a short review, the MMN is a negative component of the auditory event-related potential (ERP) that is elicited by any discriminable auditory change (“deviant”) in a sequence of repetitive (“standard”) stimuli, and it has been observed irrespective of the subject's attention or task (Näätänen, 2001; Näätänen, Paavilainen, Rinne, & Alho, 2007).

when compared to perception of speech sounds in noise or in nonnative speech perception in which cases motor brain areas seem to become essential part in decoding the complex signal (D. Callan et al., 2014; D. E. Callan et al., 2004; Du et al., 2014).

Imaging studies exploring the learning effects of new phonetic categories have also reached similar conclusions. This literature has generally reported greater task-related activation in bilateral temporal and left inferior frontal structures after compared to before training on the new learned phonetic categories (A. M. Callan et al., 2006; D. E. Callan et al., 2004; D. E. Callan, Tajima, et al., 2003; Chollet, 2000; Golestani & Zatorre, 2004; Poldrack, 2000; Van Mier, 2000). Myers and Swan (2012), for example, have observed bilateral activation in the middle frontal gyri specific to the encoding of learned category information, but detected no significant activation in the temporal lobes for between vs. within phonetic category contrasts. In this study, two groups of participants were trained to categorize speech sounds taken from a dental-retroflex-velar continuum according to two different boundary locations (one group was trained in a dental/retroflex contrast, while the other received training in a retroflex/velar contrast). After categorization training, differences in activation were observed almost exclusively in left and right frontal areas, specifically in the left inferior frontal gyrus and left and right middle frontal gyrus. The authors suggested this finding may imply that unlike the processing of novel speech categories, differential responsiveness to learned categories does not need to rely on retuning of sensitivities in the temporal lobe. This conclusion led Myers to the proposal of a model where, in the case of learned nonnative phonetic contrasts, "categorical sensitivity" emerges first in the inferior frontal lobe as participants learn the boundaries through acoustic space which define functional categories. Myers further suggests that this allows for rapid learning of category boundaries without fundamentally reshaping neural sensitivity to low-level details of the signal. (E. Myers, 2014, p. 7). Studies in which participants underwent more sustained and intensive training has uncovered similar sensitivities reflected in the posterior temporal lobe activation (Leech et al., 2009; E. B. Myers & Mesite, 2014). This model seems to be consistent with the previously proposed formulation of internal models (D. E. Callan et al., 2004; D. E. Callan, Tajima, et al., 2003; M. Iacoboni, 2008; Poeppel et al., 2008; Skipper et al., 2007; Stephen M. Wilson & Iacoboni, 2006), which was expanded by Callan et al. (2004) to model the perception of nonnative categories. In analyzing the perception of a difficult second language contrast (/l-/r/ by Japanese speakers), the authors concluded that when compared to native speakers, second language speakers revealed significantly greater activity of brain regions involved with motor planning (including PMC and Broca's area), mainly in the left hemisphere. Conversely, native speakers revealed greater activity in anterior STG/S, consistent with increased reliance on auditory-phonetic representations (D. E. Callan et al., 2004).

In summary, Myers (2014; E. B. Myers & Swan, 2012) and Callan's (D. E. Callan et al., 2004; D. E. Callan, Tajima, et al., 2003) works seem to relate well in terms of their findings and theoretical models, and both seem to be in line with other literature indicating that phonological categorization of ambiguous and second language speech sounds is thought to engage PMC processing through a dorsal pathway (Chevillet et al., 2013; Hickok & Poeppel, 2007; Lieberthal et al., 2013), where sensitivity to acoustic

information from native language speech sounds is found in the posterior superior temporal gyrus (pSTG) and superior temporal sulcus (STS) and feed-forward to category-level coding in the frontal lobe when sounds are ambiguous (E. Myers, 2014), possibly like the ones recognized as foreign language sounds.

Models of Sensorimotor Integration in Speech Processing

As extensively suggested before, the involvement of motor areas in speech perception has long been suggested in theoretical considerations of the neural processes behind speech perception (A. M. Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; A. M. Liberman & Whalen, 2000; Wernicke, 1969). The link between speech perception and production dates back to Wernicke's model of the neural network behind language processing (Wernicke, 1969), which incorporated a direct sensory-motor connection, and Liberman's 'motor theory of speech perception' (A. M. Liberman et al., 1967; Alvin M. Liberman & Mattingly, 1985; A. M. Liberman & Whalen, 2000). The motor theory assumes that speech perception is motoric in nature, i.e., the articulatory gestures are the basic units of speech perception. Categorical perception of speech is, according to this view, dependent on an innate and speech-specific module responsible for detecting the speaker's intended gestures, as opposed to being dependent on acoustic/auditory mechanisms. Although this and other controversial claims of the motor theory have since been highly questioned (Hickok, 2009; Lotto et al., 2009; Massaro & Chen, 2008; Schwartz, Basirat, Ménard, & Sato, 2012), the finding that PMC and Broca's areas are active not only during action production, but also during action observation – 'mirror system' – has ignited the discussion regarding the role of the motor system on speech perception (Arbib, 2010; Vittorio Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Perry & Bentin, 2009; Rizzolatti & Craighero, 2004; Schwartz et al., 2012).

The mirror neuron system (MNS) is especially interesting for the purposes of the current project since it has been shown that the human MNS might be the neurophysiological basis for sensorimotor integration (Galantucci, Fowler, & Turvey, 2006; V. Gallese, Gernsbacher, Heyes, Hickok, & Iacoboni, 2011). This postulation finds support in the discovery of mirror neurons in macaque area F5, thought to be a homologue of the human PMC, near Broca's area. Since these areas (PMC/Broca's area) are classically involved in speech production, a great deal of research has explored the involvement and function of the PMC and the dorsal auditory stream in speech perception (A. M. Callan et al., 2006; D. Callan et al., 2010; D. E. Callan, Callan, Honda, & Masaki, 2000; D. E. Callan, Kent, Guenther, & Vorperian, 2000; D. E. Callan et al., 2006; Meister et al., 2007; Nishitani, Schürmann, Amunts, & Hari, 2005; Stephen M. Wilson & Iacoboni, 2006). There are essentially two main classical views of perception that inspired more current theories and models exploring the involvement of the dorsal auditory stream in speech perception: a direct realist view (Gibson, 1979) and a constructivist view (Von Helmholtz, 1867).

The direct-realist theory of perception proposes that perceptual recognition is fundamentally founded on the detection of sensory stimulation properties that have parity

with self-generated actions (Fowler, 1986). In other words, gestural information is *directly* detected in speech perception, and constitutes the most basic unit of speech recognition. Direct detection, according to the direct-realist theory of speech perception, assumes no mediation by cognitive processes of inferencing or hypothesis testing. Instead, the listener directly recovers phonetically-structured articulatory gestures that are complexly encoded in the acoustic structure of the phonetic events (Fowler, 1986). The PMC, according to this view, has been indicated to implement a “non-inferential mechanism of action recognition based on neural identity” (Marco Iacoboni, 2005).

The constructivist theory of speech perception, on the other hand, ascertains that the ongoing sensory stimulation is analyzed and constrained by a process of internal stimulation that predicts sensory outcomes of articulatory gestures. This approach to speech perception is tightly related with models of visual processing in which formulated hypothesis are tested against the inherently ambiguous information available to the retina (Hatfield, 2002). Therefore, this approach does not propose a direct link between the acoustic signal and speech perception. Rather, it proposes a mediated process where hypotheses about the speech signal are formed and tested basically based on three things: sensory input, previous phonological knowledge and high-level cognitive processes (Turvey, 1974).

A more recent approach, derived from the constructivist theory, defends that the dorsal auditory stream is mainly involved with predicting sensory consequences of articulatory gestures, constraining or facilitating speech perception (D. Callan et al., 2010; M. Iacoboni, 2008; Stephen M. Wilson & Iacoboni, 2006). This “prediction” has been conceptualized as the formulation of forward internal models, i.e., mechanisms that simulate (predictively) characteristics of speech articulation and its sensory consequences (Kawato, 1999).

The concept of internal models is incorporated in a more comprehensive model of the functional anatomy of speech and language processing: the dual-stream model of speech processing (Hickok & Poeppel, 2004, 2007). This dual-stream model involves two partially segregated circuits in the brain: the ventral (occipital-temporal) and dorsal (occipital-parietal) streams (Hickok & Poeppel, 2015; Scott & Wise, 2004; Specht, 2014).

The ventral stream, also known as the ‘what’ stream (Scott & Wise, 2004), is generally agreed to be crucially involved with speech comprehension (lexical, syntactic and semantic processing) and speech perception³ (phonetic decoding, phonological and sub-lexical processing) (Hickok & Poeppel, 2007; Specht, 2013). Anatomically, the ventral stream is assumed to originate in the upper posterior part of the temporal lobe and to extend toward the anterior part of the temporal lobe, where it also connects to the ventral part of the inferior frontal gyrus through the uncinated fasciculus and extreme capsule (Saur et al., 2008; Weiller et al., 2011). It should be noted that Rauschecker and

³ However, speech perception and speech comprehension are said to be two distinct functions of the ventral stream (Davis & Johnsrude, 2007).

Scott (2009) have recently proposed a role for Broca's area and ventral inferior premotor cortex in the ventral stream, expanding the influence of motor regions in speech perception and enhancing the overlap between the two streams (in contrast with Hickok and Poeppel's (2004, 2007) dual-stream model, in which frontal speech areas are all thought to be within the postero-dorsal stream).

While the role and function of the ventral stream in speech perception and comprehension is widely understood, the role and mainly the function of the dorsal stream in that respect is still debated. Also known as the 'how' stream (Scott & Wise, 2004), the dorsal stream is believed to support auditory-motor integration, i.e., it is said to be involved in translating acoustic information of the speech signal into motor/articulatory commands (Hickok & Poeppel, 2004, 2007; Rauschecker & Scott, 2009; Scott & Wise, 2004). Anatomically, the dorsal stream extends from the posterior temporal lobe of the left hemisphere through inferior parietal areas into the left inferior frontal gyrus, also including premotor areas⁴. This stream mainly follows the arcuate fasciculus, connecting the temporal and inferior parietal lobe with the inferior frontal gyrus (Catani et al., 2007).

According to Hickok and Poeppel's (2004) proposal of the dual stream model, the two streams are assumed to be organized hierarchically, where the input of each processing step depends on the output of the previous step. Each stream is assumed to have different patterns of lateralization: the dorsal stream tends to be left lateralized, while the posterior part of the ventral stream is believed to be bilateral (Hickok & Poeppel, 2007).

Saur et al. (2008) have shown the dominating role of the dorsal stream in repetition, essentially sub-lexical repetition, like pseudo words (see also Specht, 2014). The essential role of this network in the infant development of speech, enabling the creation of motor programs through auditory-motor integration, has also been shown (Hickok, 2012; Ueno, Saito, Rogers, & Lambon Ralph, 2011). What is less understood is the actual function of the dorsal stream in speech perception and the extent to which it might impact speech comprehension (D. Callan et al., 2010; Chevillet et al., 2013; Liebenthal et al., 2013; Specht, 2014). It has been conjectured that PMC and Broca's areas may be responsible for instantiating internal models involved in constraining or aiding speech perception, especially in circumstances where the signal is degraded (D. Callan et al., 2010; D. E. Callan et al., 2004). Forward internal models, as mentioned above, predict the sensory consequences of internally stimulated articulatory models in way to facilitate speech perception. Callan and colleagues (2010; 2004) have provided some evidence that, especially in situations where the acoustic signal is degraded by noise or in nonnative speech perception, the mechanisms of internal models are important in enhancing perception "through the competitive selection of the internal model that best matches the ongoing auditory signal" (D. Callan et al., 2014, p. 2). More specifically in the case of perception of a nonnative speech contrast, in which an auditory representation

⁴ Rauschecker and Scott (2009) have specified the superior and dorsal premotor cortex as the premotor regions being influenced by the dorsal stream.

is not well defined, Callan and colleagues (2004) predicted that internal models of speech articulation would be used to facilitate phoneme identification. Indeed, their findings in a study using event-related fMRI, looking at the discrimination of /r/ vs. /l/, showed that second language⁵ speakers utilized articulatory-auditory and articulatory-orosensory based internal models to a greater extent than native speakers, with greater activation of brain areas involved with speech production planning, including the PMC (D. E. Callan et al., 2004; for similar results see also D. E. Callan, Tajima, et al., 2003; Pillai et al., 2003; Wang et al., 2003). What this study didn't use was a passive speech perception task, which would have helped ruling out potential confounds of the PMC activation related with task demands, like attentional and/or working memory resources. Furthermore, fMRI based studies cannot inform us on the timeline of the neural activation in relation to the stimulus presentation, which would inform us on when the predictive mechanisms of the internal models of speech perception become active.

On a later study, Callan et al. (2010) have somewhat extended their previous findings combining fMRI and MEG techniques to look into premotor cortex activation in relation to perceptual performance (correct and incorrect trials) of a native speech contrast. This time, their experiment used passive and active perception tasks, as well as time-frequency analysis. This study, however, did not explore differences between native and nonnative perception, which are going to be examined in the current project. Callan's findings showed greater PMC activation for correct over incorrect trials and greater event-related potentials for correct over incorrect trials in the alpha, beta and gamma frequencies prior to and after stimulus presentation. They interpreted their results as being indicative of articulatory processes facilitating perceptual performance instead of simply being a product of covert production of perceived actions. Furthermore, they have also found activation of speech motor areas in both the active and passive perception tasks, raising the possibility of articulatory processes to be necessary for speech perception (D. Callan et al., 2010).

The timing of neural activation within the dorsal auditory stream pathway, in relation to its possible role in the formulation of internal models/predictive coding in speech perception is of special interest to this current work. Constructivist approaches, based on the notion of predictive internal models, predict the occurrence of sensorimotor activation prior to and following the presentation of speech stimuli. Furthermore, it predicts greater PMC engagement in predictive coding of more difficult tasks. The current project, however, intends to disentangle working memory load processes from those related with predictive phonological coding by examining the dynamics of sensorimotor activation across the time-course of speech processing.

⁵ Japanese native speakers with some English experience (at least 6 years of classes in junior and senior high school).

EEG and Sensorimotor Integration

A large body of literature has appointed the dorsal auditory stream as an important player in the neural processes involving speech perception. It is believed that brain areas within this stream are an integral part of the sensorimotor integration processes, wherein the PMC is involved with predicting sensory consequences of internally stimulated articulatory models, consequently aiding or constraining speech perception, in line with the specifications of forward internal models (D. Callan et al., 2010; Hickok & Poeppel, 2015). Based on this type of models, important assumptions are made regarding the time-course of activation of the different brain areas involved in speech perception and also highly significant implications can be made regarding when those areas become active in relation to stimulus presentation (see D. Callan et al., 2010). As shown previously, information regarding the time-course activation of the dorsal auditory stream during perception of native and nonnative speech stimuli may expand current knowledge about the function of sensorimotor integration processes in speech perception. Research studies have reported greater sensorimotor activity in perception of nonnative speech sounds, in relation to native ones, and this activity has been revealed to increase with learning, possibly reflecting extra attentional resources necessary to decode complex speech signals (A. M. Callan et al., 2006; D. E. Callan et al., 2004; D. E. Callan, Tajima, et al., 2003; Wang et al., 2003). However, no speech event-related data is available on the perception of nonnative phonological categories. Data on this subject would provide extremely relevant information on when internal models become active during the perception of nonnative speech sounds. It is known that, during perception of native phonemes, internal models become active shortly after stimuli onset and that they can be predictive of performance (D. Callan et al., 2010). In what regards nonnative speech perception, however, we have still to uncover how these internal mechanisms function.

As described before, studies using MEG and EEG have identified oscillations in alpha and beta rhythms, generated in sensorimotor and auditory areas, to be modulated by different speech perception tasks and performance (A. Bowers et al., 2013; A. L. Bowers et al., 2014; D. Callan et al., 2010; Jenson et al., 2014; Jenson et al., 2015), and allied with new data decomposition methods it has great potential to explore the rich load of information behind event-related brain dynamics. EEG and MEG techniques have excellent time resolution (<1ms), contrary to the poor time resolution of fMRI (>1sec), for example. Hence, this section intends to describe a new method of EEG data analysis which minimizes some of the problems traditionally found with this technique, like low spatial resolution, and which provides a more adequate model of the spatially distributed event-related EEG dynamics that support cognitive events.

EEG has traditionally been considered a low-spatial resolution measure of neural activity, mainly due to a physical problem of volume conduction. For the purposes of electrophysiology, volume conduction is described as the transmission of electric or magnetic fields from an electric primary current source through biological tissue towards measurement sensors on the scalp. This causes a significant problem in the interpretation of EEG data, because signals recorded at one sensor or electrode may be influenced by more than one current source, creating the illusion of entrainment when volume

conduction might also account for changes in oscillation. For this reason, in the past two decades, the bulk of research on speech perception has been recurring to different methods of data collection and analysis, like fMRI and TMS. However, more recent shifts in thought regarding the analysis of EEG data have allowed for advances in formulas/techniques of analysis and consequently have made the investigation of temporally independent and spatially fixed source generator's feasible (Makeig, Debener, et al., 2004).

Independent Component Analysis (ICA)

A relatively novel approach to EEG data analysis applies a blind method for separation of temporally independent neural oscillations, via independent component analysis (ICA) (Makeig, Debener, et al., 2004; Makeig, Delorme, et al., 2004; Onton & Makeig, 2006; Onton et al., 2006). This method of linear data decomposition separates multichannel EEG data into maximally independent temporal and spatially fixed components with dipolar scalp maps, without taking into consideration any information about head geometry or electrode locations, strongly suggesting that recovered independent components represent physiologically distinct brain processes. Therefore, traditional EEG analysis problems like the biophysical inverse problem are bypassed by ICA. Besides, ICA is able to separate the contributions of stereotyped non-brain artifact signals, like eye movements, muscle activity, line noise, etc. (A. Delorme & Makeig, 2004; Jung et al., 2000).

In summary, analyzing EEG data using ICA presents several advantages in comparison to other traditional methods of electrophysiological data analysis. Namely, ICA is a robust method for artifact removal from EEG signal, since noise sources from blink and other artifacts are not temporally related to the cortical activity. Also, volume conduction from multiple sources and the related inversion problem are minimized because independent components are considered maximally independent from one another. Furthermore, this enhances the plausibility of a tight link between the components' oscillatory dynamics and behavior. Finally, although ICA does not present a direct solution for the source localization problem, Congedo et al. (2008) have observed low variance between dipolar source reconstructions and ICA projections for both low and high-density electrode arrays, suggesting that ICA may act as an efficient spatial filter for spatial localization.

This new approach to EEG data analysis presents, therefore, several advantages to enhance our understanding of the brain processes behind speech perception. For the purposes of the current study, an ICA decomposition of event-related EEG during speech processing tasks (perception of native and nonnative phonological contrasts) is expected to separate sources with topographic scalp maps and source reconstructions consistent with sensorimotor integration brain regions (specifically BA 1, BA 2, BA 3, BA 4, and BA 6). Of special interest to the present work is the analysis of EEG oscillations of alpha and beta components of the μ -rhythm, time-locked to stimulus events (prior to, during

and after stimulus presentation), since this rhythm is said to represent a ‘down-stream’ measure of motor activity emerging from the PMC (Jaime A. Pineda, 2005).

The Arceu or μ -Rhythm

The arceu rhythm derives its name from an arc like shape, composed of two main frequency components: one spectral peak around 10 Hz (usually within the 8 to 13 Hz band) and the other around 20 Hz (usually within the 15 to 25 Hz band) (Hari, 2006; Hari, Salmelin, Mäkelä, Salenius, & Helle, 1997; Jaime A. Pineda, 2005). The μ -rhythm is typically localized to sensorimotor regions, and its changes in oscillations are usually quite brief, ranging from 0.5 to 2 sec. (Hari, 2006; Jaime A. Pineda, 2005). While the alpha component frequencies tend to be originate from the somatosensory cortex, the beta component frequencies tend to emerge in a somatotopic manner from the precentral gyrus corresponding with the motor cortex for the effector involved (i.e., lip vs. hand movements) (Hari, 2006; Jensen et al., 2005; Jaime A. Pineda, 2005). The rhythm is said to represent a ‘down-stream’ measure of motor activity emerging from the PMC (Jaime A. Pineda, 2005).

Although the μ -rhythm can be measured over the sensorimotor cortex in the absence of processing, conceived as an ‘idling’ or ‘nil-work’ state, more complex accounts of its function are currently dominant, relating it to diverse brain functions that comprise sensory, motor, cognitive, mnemonic and integrative processes (Basar, Basar-Eroglu, Karakas, & Schürmann, 1999; Başar, Schürmann, Başar-Eroglu, & Karakaş, 1997; Petsche, Kaplan, von Stein, & Filz, 1997). Power in the μ -rhythm has been observed to decrease in association with certain tasks and to enhance with others (Karrasch, Krause, Laine, Lang, & Lehto, 1998; W. Klimesch, Doppelmayr, Russegger, Pachinger, & Schwaiger, 1998; Krause et al., 2000).

Event-Related Synchronization and Desynchronization

It is generally accepted that EEG synchronization is a correlate of brain ‘idling state’ or deactivated cortex, resulting in high spectral power and low brain metabolism. EEG desynchronization, on the other hand, is said to result from thalamocortical stimulation and it is correlated with excited or activated cortical areas, resulting in a decrease in spectral power (Jaime A. Pineda, 2005). A negative correlation has been shown to exist between the EEG and fMRI (blood oxygen level dependent (BOLD)) signal change, dependent upon the spatial scale and the frequency band of cortical synchrony, i.e., alpha-band synchrony results in high EEG spectral power over alpha rhythms, but the low overall metabolism results in a small effect on the BOLD signal in the correspondent brain regions (Formaggio et al., 2008; Gonçalves et al., 2006; Laufs et al., 2003; Yang, Liu, & He, 2010).

In what respects specifically to sensorimotor activity, there is extensive evidence relating sensorimotor event-related desynchronization (ERD) – power decrease or

suppression – with movement observation, imagination and execution (Pfurtscheller & Lopes da Silva, 1999). ERD has also been shown to correlate with task complexity and more efficient task performance (Doppelmayr, Klimesch, Hodlmoser, Sauseng, & Gruber, 2005; Wolfgang Klimesch, 1999; W. Klimesch, Doppelmayr, & Hanslmayr, 2006; W. Klimesch et al., 1996; Serman, Kaiser, & Veigel, 1996), suggesting that μ -rhythms can be task specific. Also, explicit learning of movement sequences seems to be accompanied by greater μ ERD over contralateral regions, and reduced after the movement becomes more “automatic” (Houdayer et al., 2016; Zhuang et al., 1997). This strongly suggests that activity in sensorimotor regions increases in association with learning and decreases after the task is learned. This seems to be consistent with fMRI literature reporting greater recruitment of bilateral cortical regions, including Broca’s area and premotor cortex, in association with learning of new phonological categories (D. E. Callan, Tajima, et al., 2003; E. Myers, 2014).

The relationship between sensorimotor ERD and movement observation, imagination and execution (Pfurtscheller & Lopes da Silva, 1999) strongly implicates a relationship between the function of the μ -rhythm and alpha ERD and the conceptions of the mirror neuron system (MNS), prompted by the discovery of sensorimotor neurons in the macaque F5 area, and homologue of the human PMC (Hari, 2006; Jaime A. Pineda, 2005). In fact, the possible function of μ and alpha ERD in transforming incoming sensory information into action plans has been proposed in the last two decades by several researchers (Başar et al., 1997; Graimann & Pfurtscheller, 2006; Pfurtscheller, Stancák Jr, & Neuper, 1996). Several studies have measured μ -rhythms as a function of perceived biological movements and have observed that the μ -rhythm is reliably blocked in normal subjects during the observation of a diverse range of movements, including hand, arm, mouth an even implied movement (e.g., point-light biological motion) (Crawcour, Bowers, Harkrider, & Saltuklaroglu, 2009; Muthukumaraswamy & Johnson, 2004a, 2004b; Muthukumaraswamy, Johnson, Gaetz, & Cheyne, 2006; Muthukumaraswamy, Johnson, & McNair, 2004; Ulloa & Pineda, 2007). Furthermore, μ suppression (ERD) has also been observed, although less reliably, in individuals thought to have deficits in the MNS, which suggests a functional role of μ in movement processing (Oberman et al., 2005; Oberman et al., 2013; Oberman, Pineda, & Ramachandran, 2007; Oberman, Ramachandran, & Pineda, 2008).

The opposite phenomenon to ERD is an event-related synchronization (ERS), i.e., an increase in a spectral peak at a given frequency component, i.e., an amplitude enhancement, based on the cooperative or synchronized behavior of a large number of neurons (Pfurtscheller & Lopes da Silva, 1999). This coherent activity in the alpha band is correlated with a deactivated state of the neural network, and active processing of information becomes very unlikely. μ ERS is usually interpreted as having inhibitory effects which can be very important to optimize energy demands and control over excitatory processes (W. Klimesch et al., 1996) by blocking, for example, irrelevant features in the selection of a certain phonemic category.

In summary, the reactivity of the μ -rhythm provides an easily accessible window into the complex connection between perception and action, or within the speech

perception field, the link between auditory and motor networks. The analysis of ongoing changes in the EEG signal, by means of the μ ERS and ERD, offers a unique method to enhance our understanding of the neural mechanisms that support speech perception in general, and the categorical perception of native and nonnative speech sounds in particular.

CHAPTER 3. METHODS AND PROCEDURES

Methods

Materials

The stimuli are monosyllabic pairs of English native and English nonnative sounds. The native stimuli used are /f/ and /v/, followed by the central vowel /ʌ/, forming the monosyllables /fʌ/ and /vʌ/. The nonnative stimuli are /x/ - /ɣ/, followed by the central vowel /ʌ/, forming the monosyllables /xʌ/ and /ɣʌ/. The first contrast is a minimal pair in English, and it contrasts the voiced and voiceless labiodental fricatives /vʌ/ and /fʌ/. The second contrast is nonnative to English, and it contrasts the voiced and voiceless velar fricatives /ɣʌ/ and /xʌ/.

The stimuli were all produced and recorded by a phonetically trained American English native speaker. The speaker produced thirty samples of each monosyllable at a regular speech rate and level, and maintaining a similar intonation. The best exemplars of each monosyllable, judged by similarities in overall duration, intonation curve, and vowel quality, were then selected to be used in the categorical discrimination task. The best exemplars were so judged by two linguistically trained listeners. The stimuli's intensity and duration were then normalized using PRAAT (Boersma & Weenink, 2011). The intensity of each monosyllable was manipulated to be approximately 72dB, and the duration was set to be exactly 600ms for each monosyllable.

Pairs of English native labiodental fricatives (voiced vs. voiceless), and pairs of the nonnative velar fricatives (voiced vs. voiceless) were then created and a 250ms of silence was inserted as interstimulus interval (ISI). A silence of 550ms was also inserted after each pair of monosyllables, so that the total duration of each trial would be exactly 2sec (see **Figure 3-1**).

A pilot discrimination test was implemented to make sure the nonnative stimuli were highly discriminable for English native speakers, even if recognized as not belonging to their native language. The pair of nonnative monosyllables were presented to 11 English native speakers and they were instructed to identify the sounds as being the same or different. They listened to the four possible combinations of the nonnative syllables, each repeated 10 times, resulting in a total of 40 stimuli. Mean accuracy was 93.6%, ranging from 82.5% to 100%.

Participants

One single group of 24 female monolingual American English native speakers participated in this study, with a mean age of 25.3 (range 20-69). As assessed by the

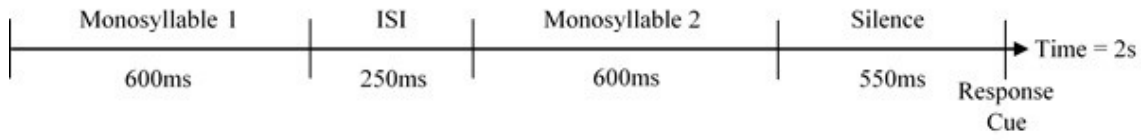


Figure 3-1. Trial timeline.

Edinburgh Handedness inventory, 23 participants were right handed and 1 was ambidextrous. All participants reported having no history of cognitive, language or hearing impairments. Prior to the experiment, informed consent approved by the University of Tennessee – Knoxville, Center of Institutional Review Board was obtained for all participants.

Procedures

The experiment was conducted in an electronically and magnetically shielded, double-walled, sound-treated booth. Participants sat in a comfortable reclining armchair with their heads and necks well supported. Participants heard audio from three different conditions: white noise, pairs of native syllables, and pairs of non-native syllables. The syllable stimuli were presented in four possible combinations, forming the following contrasts:

1. Native contrasts:

a. /fΛ/ – /vΛ/; /vΛ/ – /fΛ/; /fΛ/ – /fΛ/; /vΛ/ – /vΛ/.

2. Nonnative contrasts:

a. /xΛ/ – /γΛ/; /γΛ/ – /xΛ/; /xΛ/ – /xΛ/; /γΛ/ – /γΛ/.

Stimuli were presented using Stim 2 4.3.3 stimulus presentation software on a PC computer. Participants were always presented with the passive listening conditions (native and nonnative) first, so that no tendency to covertly categorize the sounds would carry over. In the active tasks, participants were instructed to actively discriminate the above described native and nonnative syllable pairs by pressing a button. It has been shown that premotor planning occurs in repeated button-press movements 1 second prior to muscle contraction (Grimm & Pfurtscheller, 2006; Hari, 2006) and sensorimotor suppression (ERD) peaks shortly after (200 ms) (Makeig, Delorme, et al., 2004). In order to control for the possibility that preparation for the response might confound motor activity related to stimulus processing, participants' manual response was cued by a 100 ms, 1000 Hz, sawtooth wave tone, which appeared 2000 ms after stimulus onset. The passive listening conditions were also followed by button-press, in order to control for the required button-press in the active conditions (see below). The button-press in the passive listening conditions also serves to ensure participants are paying attention to the stimuli presented. Furthermore, button-press order (right or left hand) was counterbalanced across participants and experimental conditions. The presence of a button press in all conditions ensures that any differential activity is not a result of motor activity related to button press motion, but instead a result of the demands of the different conditions.

Participants were, in summary, asked to listen under the following conditions:

1. Passively listening to noise (Pas);

2. Passively listening to native syllables (PN);
3. Passively listening to non-native syllables (PNN);
4. Active native syllable discrimination (ND);
5. Active nonnative syllable discrimination (NND).

The order in which the conditions were presented was pseudo-randomized: the passive conditions – PN and PNN – were presented first and counterbalanced across participants, then the active conditions – ND and NND – and the noise condition – Pas – were presented in a randomized order. All conditions were presented in two blocks consisting of 40 trials each, resulting in a total of 560 trials. All stimuli were presented at an absolute intensity of 72 dB. An example timeline of one stimulus trial is displayed below, in **Figure 3-2**. To present the audio stimuli, audio was routed through insert ER-1-14A ear tips. Electroencephalogram (EEG) data was acquired and analyzed using a Neuroscan 4.3.3 system. The total duration of EEG data collection was approximately 60 minutes.

Data Acquisition

Methods applied to EEG data acquisition follow methods employed in Bowers et al. (2013) and Jenson et al. (2014).

Sixty-eight electrode channels were used to acquire EEG data based on the extended international 10–20 method of electrode placement (Klem, Luders, Jasper, & Elger, 1999) using an unlinked, sintered NeuroScan Quik Cap. All recording electrodes were referenced to the common linked left (M1) and right (M2) mastoids. The electro-oculogram (EOG) was recorded by placing electrodes on the left superior orbit and the left inferior orbit (VEOG) as well as the lateral and medial canthi of the left eye (HEOG) to monitor vertical and horizontal eye movements, respectively.

EEG data were collected using Compumedics NeuroScan 4.3.3 software and Synamps 2 system. The raw data EEG data were filtered (0.15–100Hz) and digitized via a 24-bit analog-to-digital converter at a sampling rate of 500Hz. Data was time-locked to the onset of acoustic stimuli (time = 0).

Data Processing

Following data collection, the recorded EEG data were processed using EEGLAB 12 software (A. Delorme & Makeig, 2004). The several steps involved in individual and group data processing and analysis are itemized below.

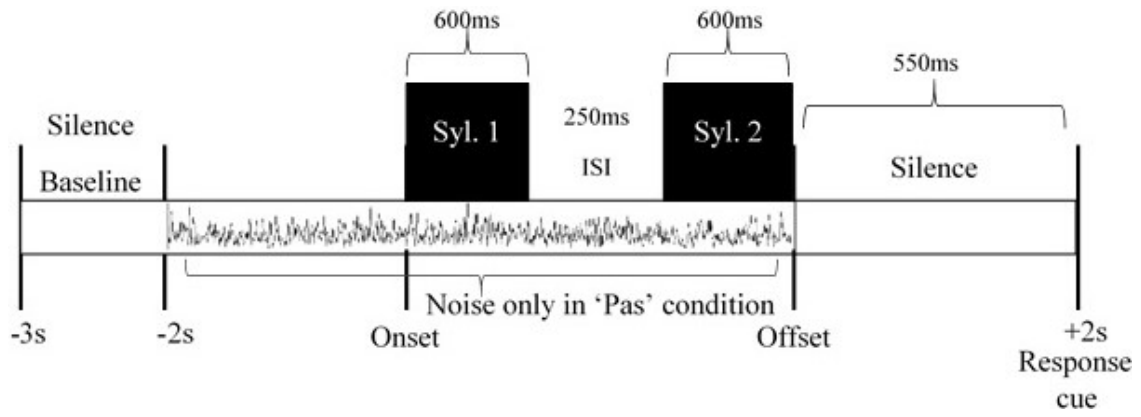


Figure 3-2. Stimuli presentation timeline.

A. Steps for individual processing:

1. Processing of 10 raw EEG files (5 conditions x 2 blocks) for each participant;
2. independent component analysis (ICA) on all concatenated files across all conditions for each participant;
3. localization of neural and non-neural dipoles and identification of independent components (IC).

B. Steps for group processing:

1. analysis of the neural ICs, using the STUDY module of EEGLAB 12;
2. identification and cluster common components across participants by the application of principal component analysis (PCA);
3. identification of μ clusters (left and right) from the neural STUDY;
4. localization of μ clusters, using equivalent current dipole (ECD) analyses;
5. time-frequency analyses (changes in spectral power across time), by measuring event-related spectral perturbations (ERSPs) of the μ clusters (left and right).

The details of the analysis on each step are described next.

Individual EEG data processing/analyses

For each condition, raw data from the two blocks of 40 trials each were:

1. appended to make a single 80 trial data set for each condition;
2. down sampled to 256Hz, in order to decrease computational requirements for ICA processing;
3. epoched into 5000ms segments with individual epochs spanning from -3000 to $+2000$ ms around time zero;
4. bandpass filtered between 3-34Hz to ensure that alpha and beta could be identified while filtering muscle movement from surrounding frequencies;
5. re-referenced to mastoid electrodes;

6. visually inspected for gross artifacts, which were manually removed together with incorrect responses or response latencies that were abnormally long when compared to the mean latency for each participant. Only correct trials were included in the following analysis to ensure participants were attending to the stimuli and task in hand.

A minimum contribution of 40 epochs per participant and condition was required for inclusion in the experiment. The average number of usable trials across participants per condition, however, exceeded the minimum of 40 required for inclusion in the experiment.

After following the above described steps to pre-process the individual data, and prior to ICA training, the pre-processed EEG data for each individual participant were concatenated across conditions, in order to obtain one single set of ICA weights. This allowed for an activity comparison across conditions within spatially fixed ICs, i.e., all conditions for each subject bear the same set of component weights, which allows for a fair comparison between conditions. An extended Infomax algorithm (T. W. Lee, Girolami, & Sejnowski, 1999) was then used to decorrelate the data matrix, prior to ICA rotation. The data matrix was subsequently subjected to ICA training using the ‘extended *runica*’ algorithm in EEGLAB 12. The initial learning rate was set to 0.001 and stopping weight of 10⁻⁷. After decomposition, 66 ICs were yielded for each participant, reflecting the total number of recording electrodes (68 – 2 reference electrodes). Scalp maps for each IC were obtained by projecting the inverse weight matrix (W^{-1}) back onto the spatial EEG channel configuration.

After ICA decomposition, ECD models for each component were computed, using a standard 4-shell spherical model in the DIPFIT toolbox (Oostenveld & Oostendorp, 2002), freely available at <https://sccn.ucsd.edu/eeglab/>. Standard 10-20 electrode coordinates were warped to the head model followed by coarse and fine-fitting to the spherical wire matrix, yielding a single dipole model for each 1584 ICs (66 ICs x 24 participants). Dipole localization requires back-projecting the signal to a source that may have generated the scalp distribution for a given IC, and then computing the best forward model to explain the highest percentage of scalp map variance (Arnaud Delorme, Palmer, Onton, Oostenveld, & Makeig, 2012). The residual variance (RV) in dipole localizations were also computed, referring to the potential mismatch between the initial scalp map and the forward projection of the ECD model.

Group EEG data processing/analyses

The EEGLAB STUDY module was used to conduct group data analyses. This module allows for the analyses of ICA data from multiple participants across conditions, using specified designs relevant for the hypothesis in test. For the purposes of the current study, several contrasts were run to analyze possible significant differences between passive and active conditions as well as native and nonnative conditions. In order to

measure neural activity, in the STUDY module options' specifications, only neural, i.e., 'in head', dipoles with a residual variance (RV) < 30% were analyzed. To be included in further analysis, however, only dipoles with a RV < 25% were considered and the average RV of the included dipoles in the left and right μ clusters ended up being < 15%.

After the STUDY analysis was complete, component pre-clustering was performed, using the K-means statistical toolbox, part of the EEGLAB toolbox (A. Delorme & Makeig, 2004), which uses information based on common scalp maps, dipoles and spectra to group similar components from each participant via PCA. The resulting neural clusters were then individually inspected in terms of their spectra, scalp maps and dipoles, with special attention payed to the clusters of interest: left and right μ clusters and neighboring clusters. Localization to BA1-4, and 6 (i.e., somatosensory regions, primary motor and premotor regions), with characteristic μ spectra were the final inclusion criteria necessary for inclusion in the μ clusters.

The μ clusters source localization was identified using the DIPFIT module of the EEGLAB toolbox. This module provides an ECD source localization based on the average coordinates (x, y, z) of all IC dipoles within each cluster. The spherical dipoles coordinates were then converted to Talairach coordinates. The source localization of each IC dipole was then retrieved by inserting the coordinates in the Talairach Client application (Lancaster et al., 1997; Lancaster et al., 2000), freely available at <http://www.talairach.org/>.

Following the identification and inspection of the μ clusters, an ERSP analyses was employed to compute changes (scaled in normalized dB units) in power across time (i.e., time-frequency analysis) within the spectral range of interest (4-33Hz). Time-frequency transforms were derived using a Morlet sinusoidal wavelet set at 3 cycles at 3Hz, rising linearly to 20 cycles at 40Hz. The present study used a 900ms pre-stimuli period, selected from -2900 to -2000ms time interval, as the baseline for each trial. This baseline was constructed from a surrogate distribution based on estimates of spectral power from 200 randomly selected latency windows from within the 1000ms inter-trial interval (Makeig, Debener, et al., 2004). Subsequent individual ERSP changes from baseline over time were computed using a bootstrap resampling method ($p < 0.025$ corrected for false discovery rate (FDR)). The single trial current for all experimental conditions for frequencies between 7 and 27Hz and times from -500 to 2000ms were entered in the time-frequency analyses.

CHAPTER 4. RESULTS

Discrimination Accuracy

Participants' average discrimination accuracy for the native contrast was 98.8% (range: 95-100%). Participants' average discrimination accuracy for the nonnative contrast was 97.3% (range: 86-100%). Although participants reported difficulties discriminating the nonnative contrast, their discrimination accuracy was still quite close to the average discrimination accuracy in the native contrast. Only correct trials were used in the EEG analysis to ensure that participants were paying attention to the stimuli and task in hand.

μ Clusters

Right and left μ -components, localized to sensorimotor regions, were found across speech perception conditions. Nineteen participants contributed with thirty-nine IC's presenting distinctive markers of the left μ -rhythm and 22 participants contributed with forty-seven IC's presenting distinctive markers of the right μ -rhythm (see **Figure 4-1** and **Figure 4-2**). The mean source location of the μ clusters was identified within the frontal lobe, precentral gyrus, Brodmann area 6, according with the following Talairach coordinates: [(x, y, z) -40, -5, 51] in the left hemisphere; [(x, y, z) 41, -6, 49] in the right hemisphere. The mean residual variance (RV) on both the left and right μ clusters was less than 9% (8.42% and 8.55% respectively).

Event-Related Spectral Perturbations

Similar patterns of results were found in both the left and right hemispheres, the left generally presenting slightly stronger power. For this reason, although always presenting figures for both the right and left hemispheres, the presentation will focus on the left μ mean ERSP's.

Left and right μ mean ERSP's across subjects and conditions are depicted in a time-frequency graph with bootstrapped ($p < 0.05$)⁶, FDR corrected significance values for condition in the far-right panel of **Figure 4-3**. The analysis revealed significant ERSP differences across conditions, with stronger alpha and beta event-related decrease in spectral power (ERD) after stimuli presentation offset (1450ms). A shorter ERD effect is also seen in the alpha band range precisely at the offset of the first syllable stimulus, around 600ms.

⁶ The FDR alpha value in EEGLAB was set to 0.025 for an effective alpha value of 0.05.

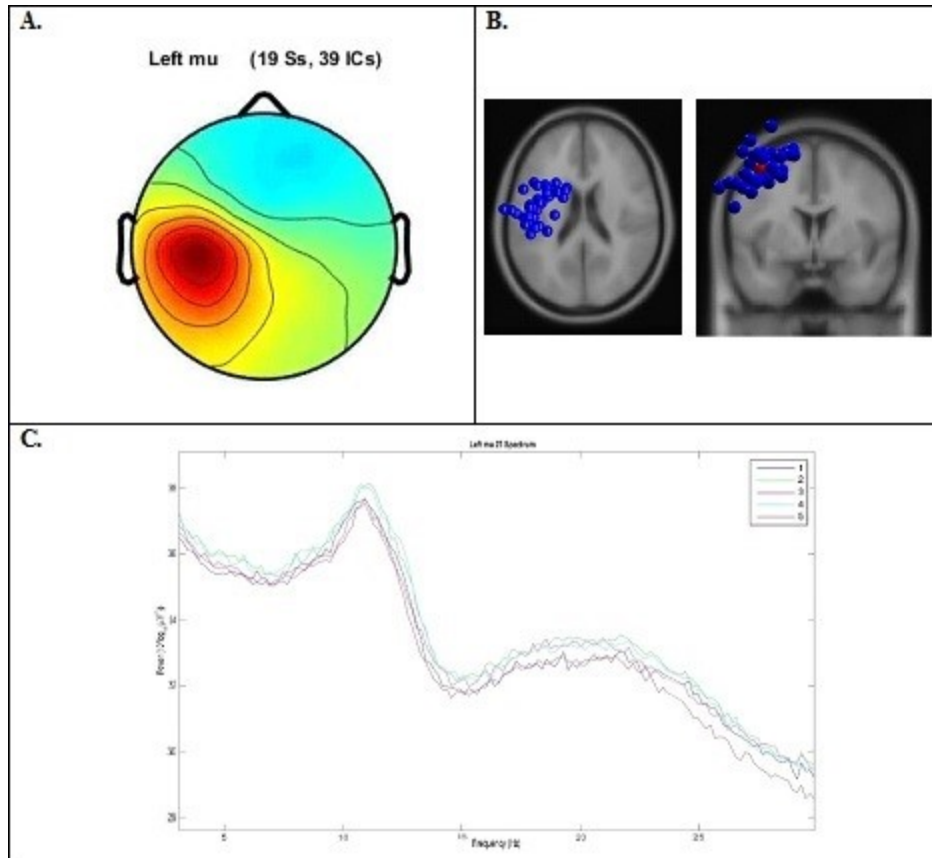


Figure 4-1. Cluster results for the left-hemisphere μ components.

A) Mean scalp potential distribution (W-1) scaled to RMS microvolts and individual scalp distributions for each participant. B) Average equivalent current dipole location. C) Mean spectra of the component as a function of condition.

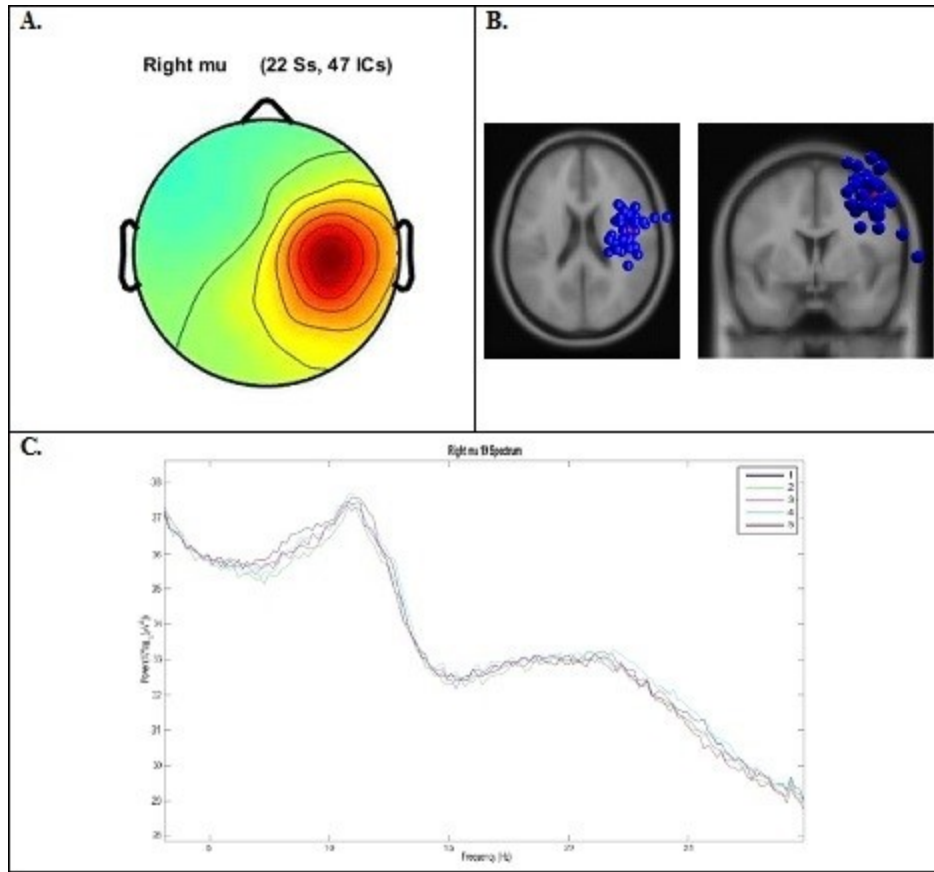
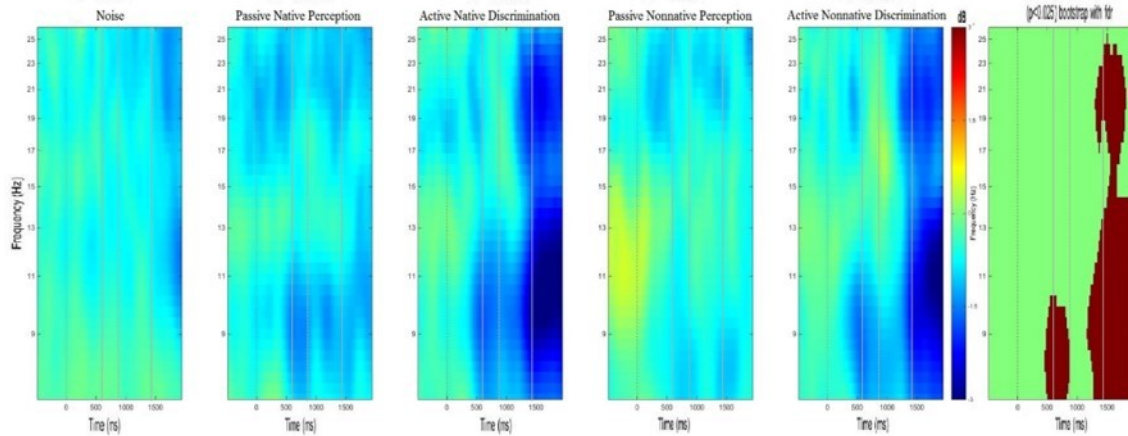


Figure 4-2. Cluster results for the right-hemisphere μ components.

A) Mean scalp potential distribution (W^{-1}) scaled to RMS microvolts and individual scalp distributions for each participant. B) Average equivalent current dipole location. C) Mean spectra of the component as a function of condition.

Left μ



Right μ

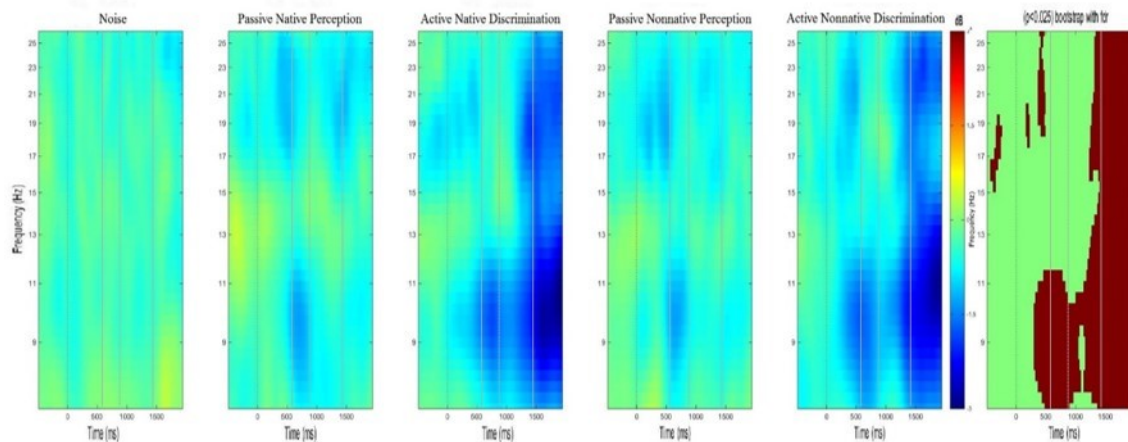


Figure 4-3. Mean left and right hemisphere μ time-frequency ERSP's. ERSP's are scaled in the same RMS decibel units as a function of condition (1x5) and random effects analysis in the traditional alpha (8-13Hz) and beta (13-25Hz) ranges. Non-significant values are colored green and significant values are colored red. Event-related decreases in spectral power are indicated in a blue to teal scale (-3dB) and increases are indicated in a yellow to red scale (3dB). Dotted line marks the syllable onset and full line marks the syllable offset.

These differences appear to be mainly driven by differences between the active and passive tasks, since alpha and beta ERD are mostly observed in the active tasks, after stimuli offset (1450ms). To determine the validity of this observation, a separate ANOVA design was computed, using the STUDY module commands, to contrast passive listening to speech stimuli with active discrimination of the syllables.

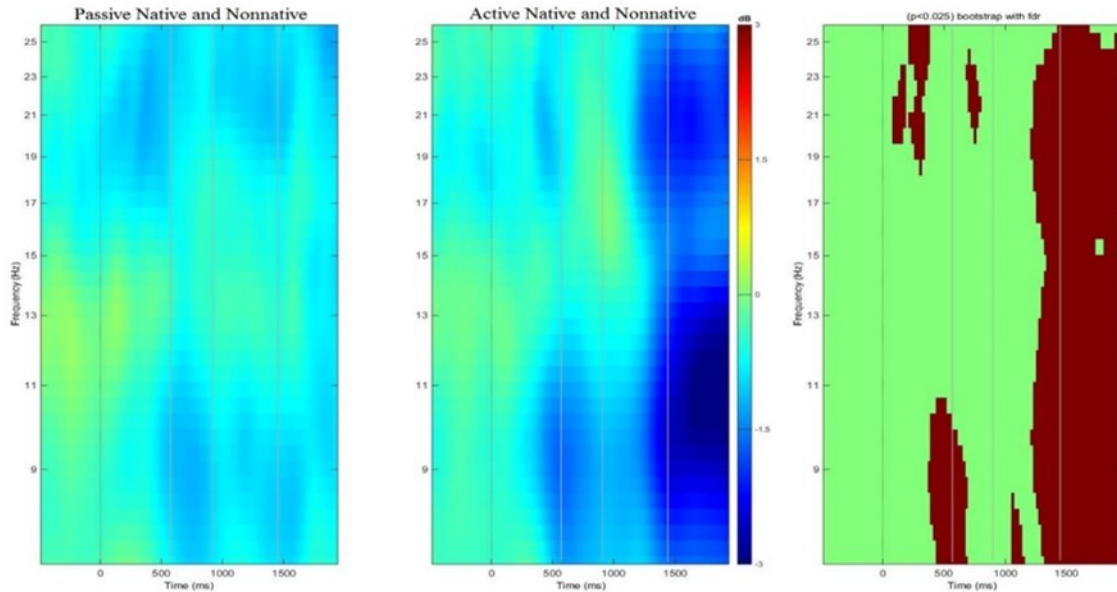
The data revealed significantly ($p^{FDR}<0.05$) higher power for μ alpha and beta ERD in active speech perception conditions, when compared to passive speech perception ones. This effect was evident both during and especially after stimuli presentation (**Figure 4-4**) with the greatest areas of activation beginning approximately 300ms after syllable offset. Although weak alpha and beta ERD is visible in the passive task, in the time between the presentation of the two syllables in the stimuli (around 600 to 800ms), its presence is significantly stronger in the active tasks.

A similar pattern of results was observed for comparison of passive and active tasks in both the native and nonnative conditions. In the native condition, significantly greater ($p^{FDR}<0.05$) alpha and beta ERD were found for the active task, when compared to passively listening to native sounds, in the timeline after stimuli offset (1450ms) (**Figure 4-5**). In the nonnative condition (**Figure 4-6**), the active discrimination of nonnative sounds was found to be related with significantly stronger alpha ERD, in the 500-600ms time range, coincident with the offset of the first syllable sound (600ms), as well as stronger alpha and beta ERD after stimuli offset (1450ms).

In order to compare sensorimotor activation during passive perception of native phonemes in comparison to nonnative phonemes, an ANOVA design was computed, using the STUDY module, to individually contrast passive native versus passive nonnative perception. **Figure 4-7** depicts significantly ($p^{FDR}<0.05$) stronger alpha (8-14Hz) and beta (22-25Hz) ERS before and during early stimuli onset in the passive nonnative task. On the other hand, the passive native perception task presents alpha and beta ERD before stimuli onset as well as stronger alpha ERD during the time between presentation of the two syllables composing the stimuli (600-850ms), and during the presentation of the second syllable (1000ms) (**Figure 4-7**).

A separate ANOVA was conducted to further explore the differences between sensorimotor activation in passive speech perception tasks and the baseline noise condition. The passive speech perception tasks in general, both for native and nonnative conditions, present significantly ($p^{FDR}<0.05$) stronger alpha and beta ERD during and between syllable presentation (600-800ms), as well as significantly stronger alpha ERD right after stimuli offset (1450ms), when compared to the noise condition (**Figure 4-8** and **Figure 4-9**). In addition, when compared to the passive speech perception task, the noise condition presents overall greater increases in power (ERS), both in alpha and beta bands, consistent with inhibition of active phonological processing (**Figure 4-8** and **Figure 4-9**). It is noteworthy, however, that the nonnative condition presents significantly greater alpha (8-14Hz) ERS before and during the first syllable presentation, compared to the noise condition.

Left μ



Right μ

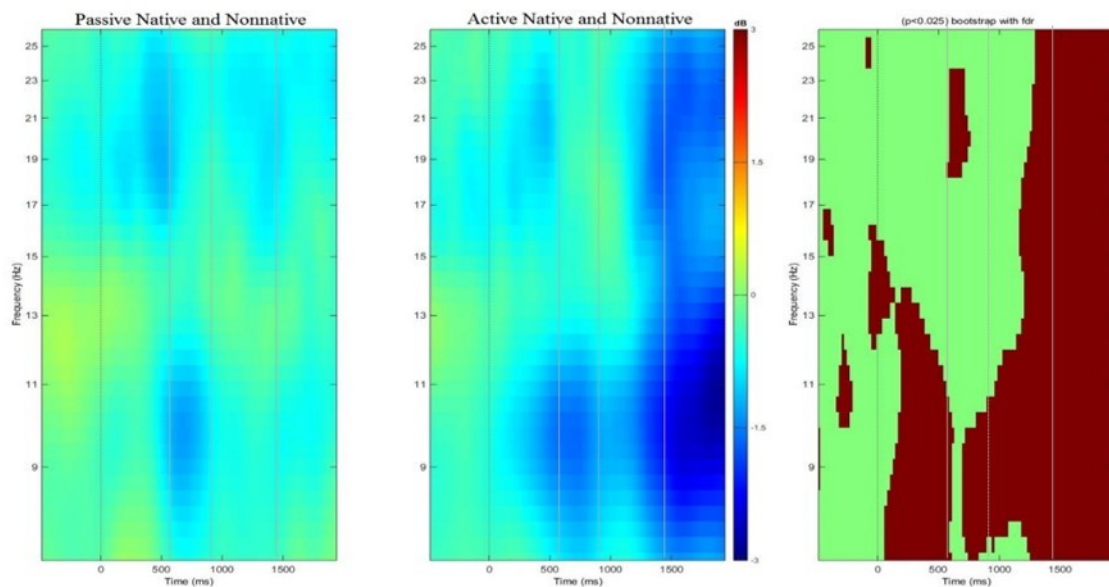
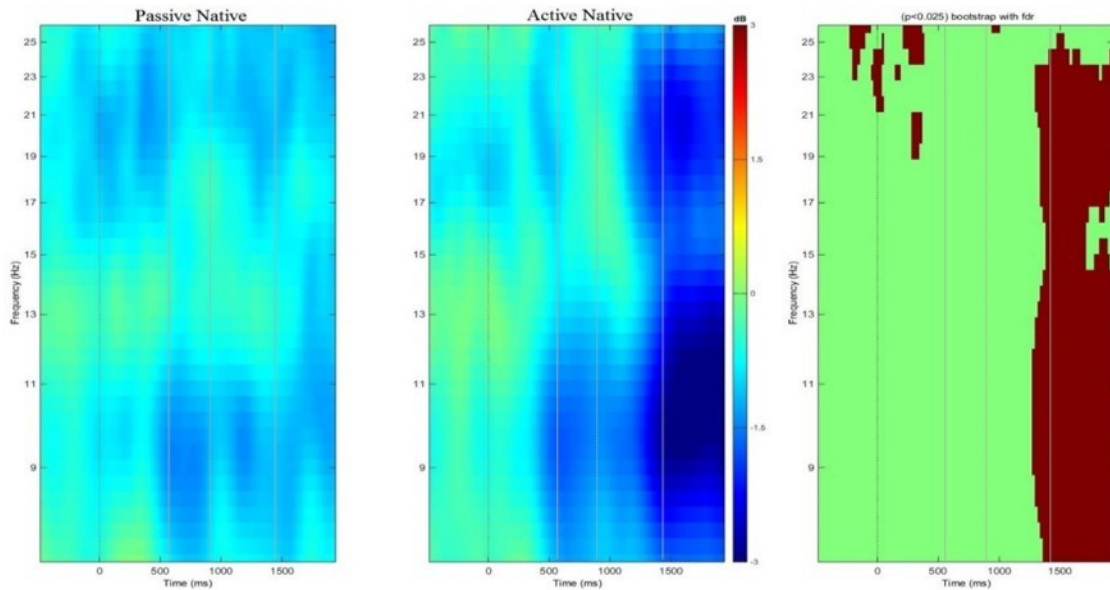


Figure 4-4. Mean left and right hemisphere ERSP's for passive (native and nonnative) versus active (native and nonnative) perception tasks. ERSP's are scaled in the same RMS decibel units as a function of condition (1x2) and random effects analysis in the traditional alpha (8-13Hz) and beta (13-25Hz) ranges. Non-significant values are colored green and significant values are colored red. Event-related decreases in spectral power are indicated in a blue to teal scale (-3dB) and increases are indicated in a yellow to red scale (3dB). Dotted line marks the syllable onset and full line marks syllable offset.

Left μ



Right μ

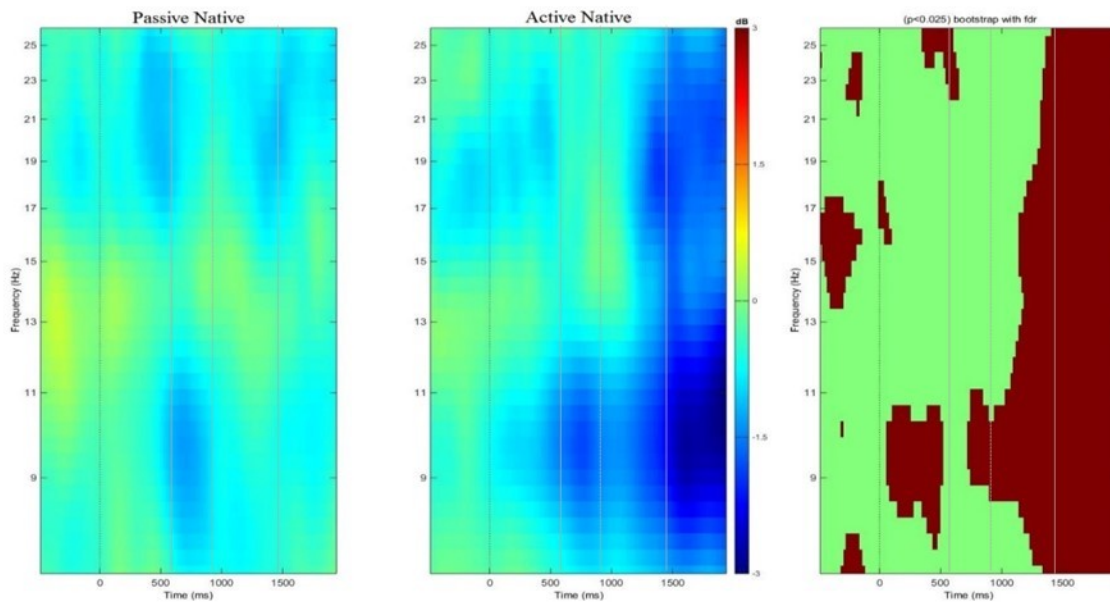
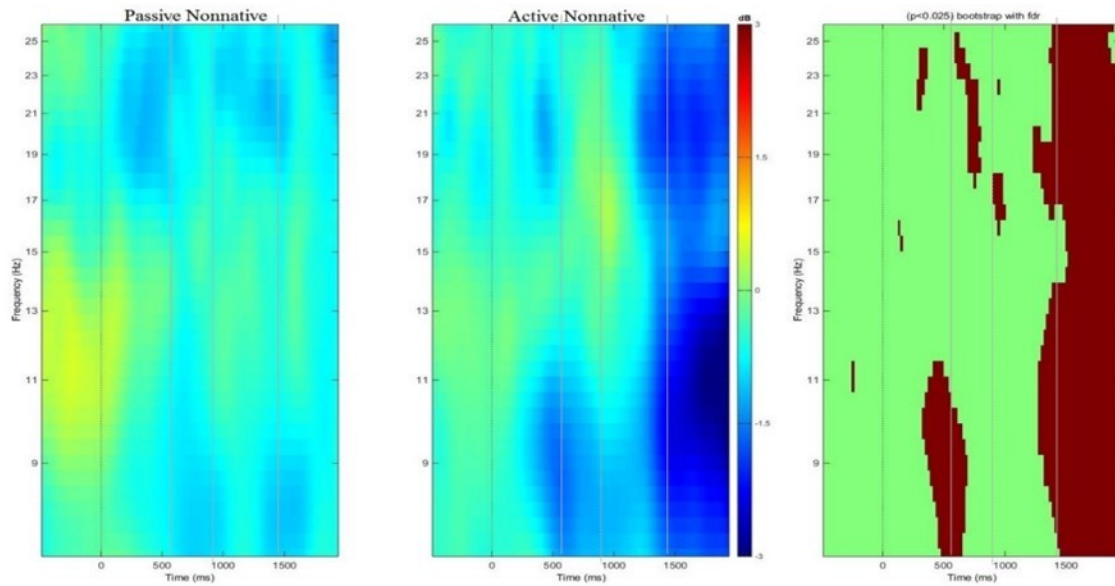


Figure 4-5. Mean left and right hemisphere ERSP's for passive native versus active native perception tasks.

ERSP's are scaled in the same RMS decibel units as a function of condition (1x2) and random effects analysis in the traditional alpha (8-13Hz) and beta (13-25Hz) ranges. Non-significant values are colored green and significant values are colored red. Event-related decreases in spectral power are indicated in a blue to teal scale (-3dB) and increases are indicated in a yellow to red scale (3dB). Dotted line marks the syllable onset and full line marks syllable offset.

Left μ



Right μ

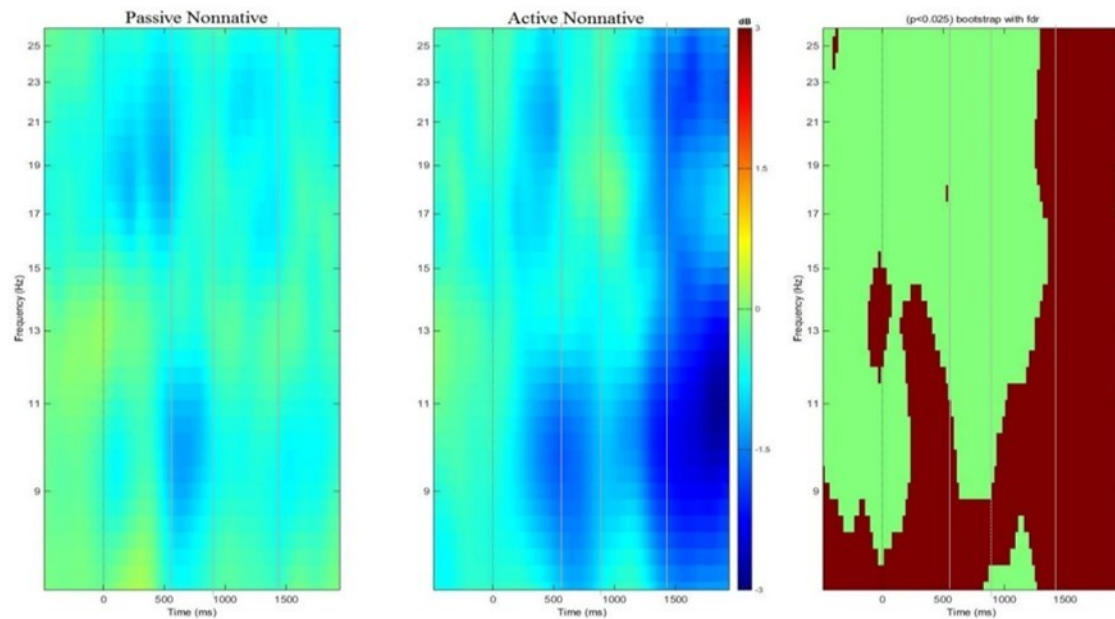


Figure 4-6. Mean left and right hemisphere ERSP's for passive nonnative versus active nonnative perception tasks.

ERSP's are scaled in the same RMS decibel units as a function of condition (1x2) and random effects analysis in the traditional alpha (8-13Hz) and beta (13-25Hz) ranges. Non-significant values are colored green and significant values are colored red. Event-related decreases in spectral power are indicated in a blue to teal scale (-3dB) and increases are indicated in a yellow to red scale (3dB). Dotted line marks the syllable onset and full line marks syllable offset.

Left μ

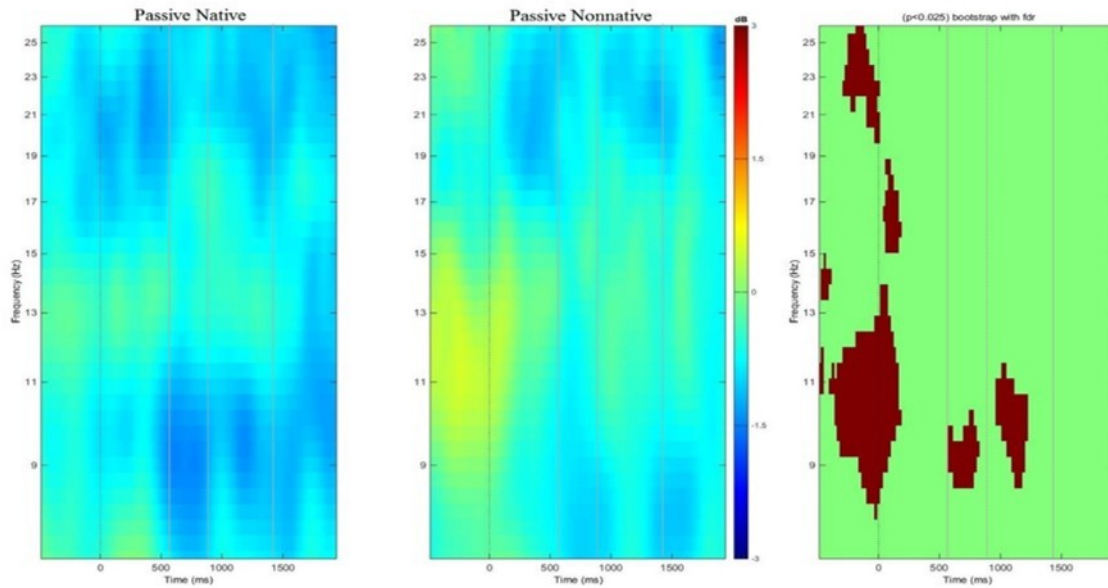
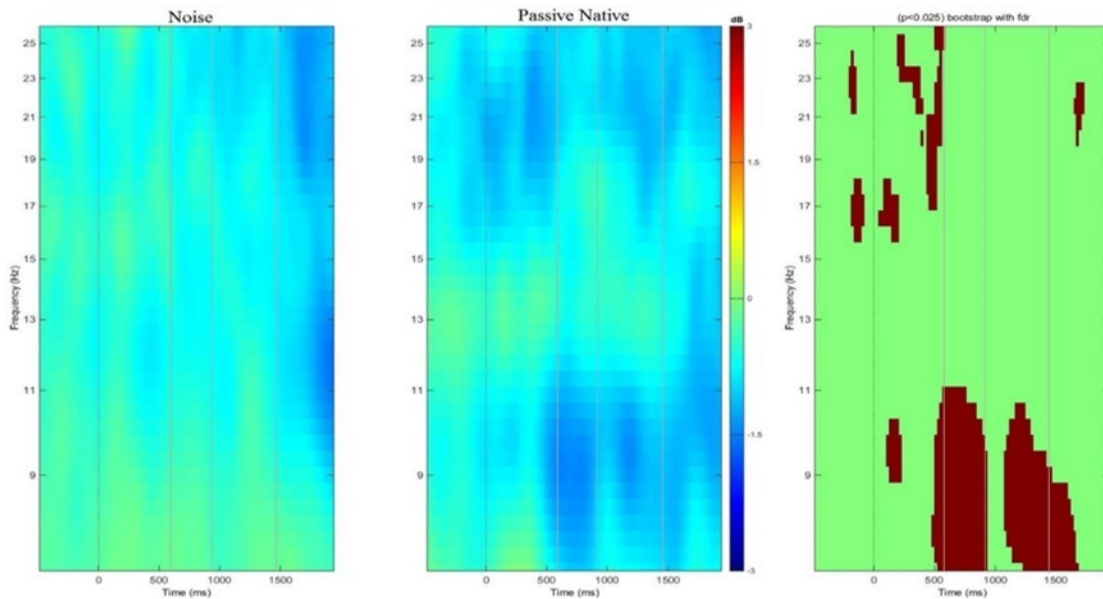


Figure 4-7. Mean left hemisphere ERSP's for passive native versus passive nonnative speech perception tasks.

ERSP's are scaled in the same RMS decibel units as a function of condition (1x2) and random effects analysis in the traditional alpha (8-13Hz) and beta (13-25Hz) ranges. Non-significant values are colored green and significant values are colored red. Event-related decreases in spectral power are indicated in a blue to teal scale (-3dB) and increases are indicated in a yellow to red scale (3dB). Dotted line marks the syllable onset and full line marks syllable offset.

Left μ



Right μ

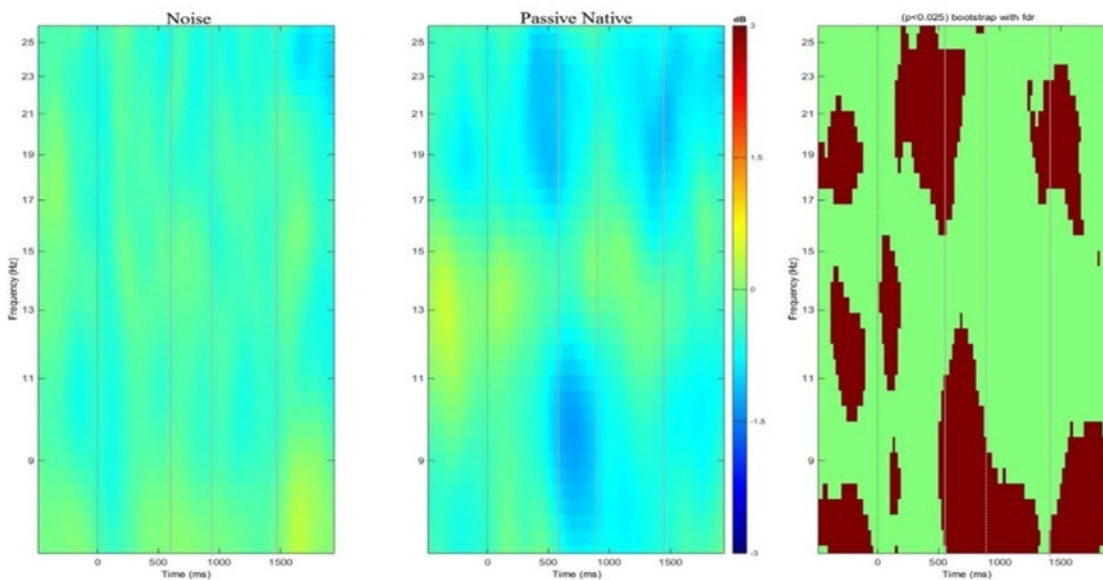
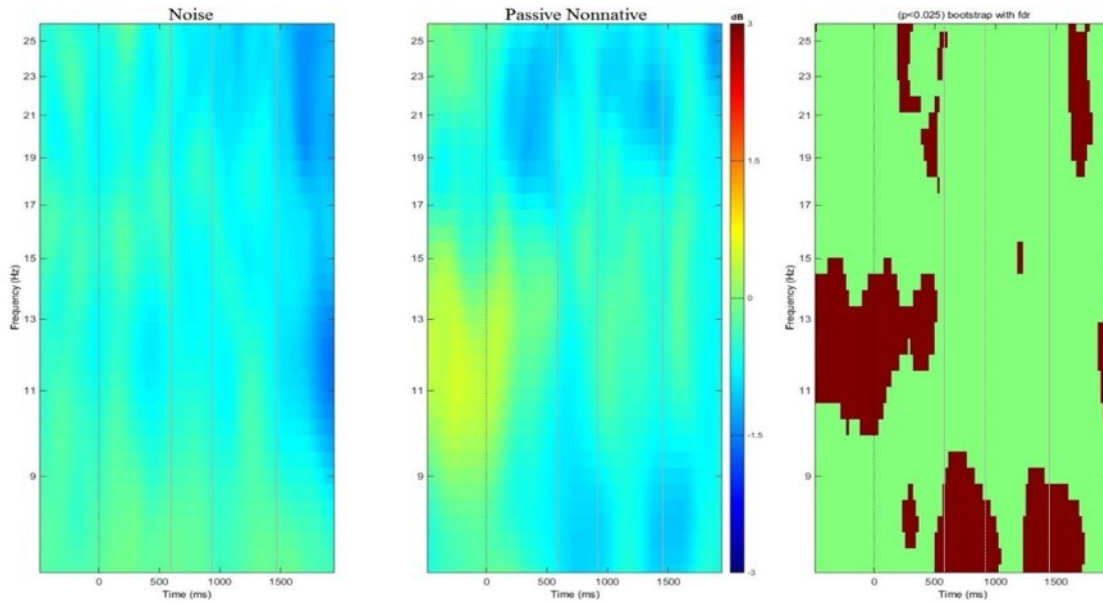


Figure 4-8. Mean left and right hemisphere ERSP's for listening to noise versus passive native speech perception task.

ERSP's are scaled in the same RMS decibel units as a function of condition (1x2) and random effects analysis in the traditional alpha (8-13Hz) and beta (13-25Hz) ranges. Non-significant values are colored green and significant values are colored red. Event-related decreases in spectral power are indicated in a blue to teal scale (-3dB) and increases are indicated in a yellow to red scale (3dB). Dotted line marks the syllable onset and full line marks syllable offset.

Left μ



Right μ

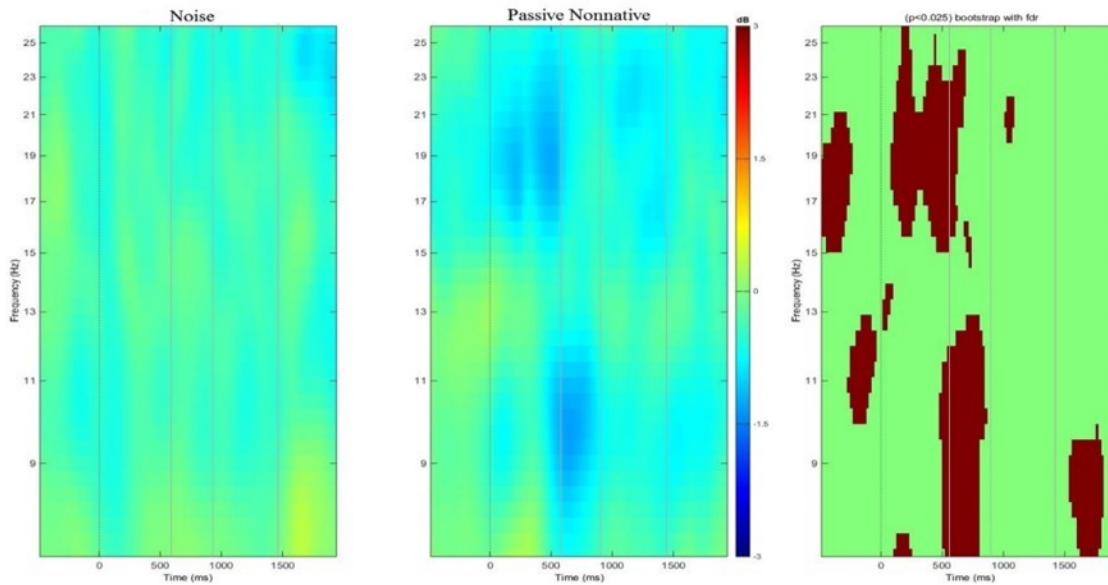


Figure 4-9. Mean left and right hemisphere ERSP's for listening to noise versus passive nonnative speech perception task.

ERSP's are scaled in the same RMS decibel units as a function of condition (1x2) and random effects analysis in the traditional alpha (8-13Hz) and beta (13-25Hz) ranges. Non-significant values are colored green and significant values are colored red. Event-related decreases in spectral power are indicated in a blue to teal scale (-3dB) and increases are indicated in a yellow to red scale (3dB). Dotted line marks the syllable onset and full line marks syllable offset.

CHAPTER 5. DISCUSSION

Experimental Results

The main aim of the current project was to disentangle working memory and automatic phonological related functions of the dorsal auditory stream during speech perception, via analysis of its time-course activity. First, it was hypothesized that activation related to automatic phonological processing would occur during stimulus presentation and that processing related to working memory would occur after stimulus presentation. In previous research, μ sensorimotor activity related to maintenance of working memory and decision making processes has been interpreted to occur shortly after stimuli offset and sustained until sensory decision is reached, not diminishing in power before a response is given (A. Bowers et al., 2013; Jenson et al., 2014), consistent with the idea of holding a mental representation in memory until a task is complete. However, the instantiation of automatic phonological processing, related to the formulation of general predictions regarding likely incoming sensory signals, should occur earlier than the memory related mechanisms. If, in fact, these internal models are associated with predictions made about the possible articulatory consequences of the auditory sensory input, these mechanisms should happen right after acoustic analysis of the speech signal in superior temporal regions, in accordance with theories based on the dual-stream model. MMN studies indicate that a time window of about 200ms to 300ms after the stimuli onset is necessary to process an auditory/acoustic speech signal (Näätänen, 2001; Zhang, Kuhl, Imada, Kotani, & Tohkura, 2005). Furthermore, consistent with this time window, in a study examining the role of the dorsal auditory stream in phonological processing, using event-related potentials and fMRI, Liebenthal and colleagues (2013) found early and almost simultaneous activity in superior temporal regions (80-100ms after stimulus onset) as well as in somatomotor brain areas (95-230ms after stimulus onset), indicating that the formulation of internal articulatory models of speech occur quite early in the neural processing of the speech signal, and may, in fact, be automatic and independent of working memory maintenance.

Having these hypotheses in mind, regarding the possible time-course of sensorimotor activation within the dorsal auditory stream, it was further hypothesized that automatic phonological processing should be present during passive speech perception tasks as well as active tasks, while working memory processes should occur prominently in active discrimination tasks. This hypothesis finds some support in research showing that tasks requiring active discrimination of phonemes represent greater dorsal stream activation than passive listening to the same phonemes (Jussi Alho et al., 2014; D. Callan et al., 2014; Meister et al., 2007). Greater cognitive demands of the task (active discrimination, discrimination in noise, etc.), related to greater effort and/or greater reliance on a phonological working memory, seem therefore to elicit greater activity in the dorsal stream. Our results are consistent with our hypotheses, showing greater sensorimotor activation in active tasks compared to passive ones, after stimuli offset. The greater sensorimotor involvement in active tasks after stimuli presentation may be related to covert rehearsal/replaying of the speech stimuli in order to maintain it in working

memory or evaluation of the stimuli in the decision-making process. These results are also congruent with findings of post-stimuli alpha ERD in other speech perception related studies (Cuellar, Bowers, Harkrider, Wilson, & Saltuklaroglu, 2012; Obleser & Weisz, 2012; J. A. Pineda et al., 2013; Shahin, Picton, & Miller, 2009).

Furthermore, the results also revealed greater sensorimotor activity during passive speech perception when compared to listening to noise. As far as we know, this is the first experiment providing data that compares μ sensorimotor activity related to speech perception in comparison to a noise signal. The results observed are consistent with previous findings indicating that the μ -rhythm is responsive to biologically relevant/reproducible sounds (S. M. Wilson et al., 2004). Although significantly weaker μ sensorimotor activity was found in passive speech tasks when compared to active tasks, it is extremely interesting to find that passively listening to speech elicits a certain level of sensorimotor activity which is not present when listening to noise. It is worth noting here that passive conditions were always presented before active ones in order to eliminate the possibility of priming the participant to discriminatory behavior. Instead, participants were simply instructed to listen to the syllables being presented and press a button after the beep, which happened 2000ms after stimuli onset. This was done to make sure that participants were minimally attending to the stimuli, given that Möttönen et al. (2014) suggested that a certain level of attention to the speech stimuli is necessary to activate sensorimotor processes. Our results are, therefore, in line with previous findings suggesting that listening to speech, irrespective of the task, automatically elicits sensorimotor processes (Murakami, Kell, Restle, Ugawa, & Ziemann, 2015; Stephen M. Wilson & Iacoboni, 2006; S. M. Wilson et al., 2004).

Our findings, however, seem to be inconsistent with those found by Bowers et al. (2013; 2014). They examined differences in sensorimotor activity, using EEG, in active vs. passive speech perception in noise and did not find significantly greater sensorimotor activity in passive speech sound perception, when compared to a noise baseline. The reason for this difference is not clear, but it is possible that degrading both signals with noise may account for the discrepancy between their findings and ours, since discrimination in noise has generally been found to relate with greater motor activity. In fact, in more recent studies by the same research group (A. L. Bowers et al., 2014; Jenson et al., 2014) it was found that alpha and beta rhythms, generated in the sensorimotor and auditory areas and measured during an EEG phoneme discrimination tasks, were stronger during discrimination in noise as compared to discrimination in quiet (both active tasks). Other researchers have also found greater PMC involvement in the discrimination of speech sounds in noise, indicating an aiding role of the PMC in decoding degraded speech signals (D. E. Callan, Jones, et al., 2003; D'Ausilio et al., 2012; Du et al., 2014). It is then possible that in the passive task, employed in Bowers (2013), participants were not attending to the speech signal at all because of the noise, i.e., the inclusion of noise in the speech signal was sufficient to block automatic phonological processing during the passive task due to the lack of effort being expended to parse the speech from the noise. That this is reasonable is again suggested by Möttönen et al. (2014), given that her findings point to the necessity of a certain level of attention to the speech stimuli in order to activate sensorimotor processes. In this study it was found that when speech sounds

were attended (active perception task) the articulatory-motor cortex contributed to the auditory processing of the sounds earlier (at 60-100ms after stimulus onset) than when the speech sounds were unattended (passive perception task) (~170ms after stimulus onset). Thus, although sensory-motor activation may occur automatically when a participant hears speech, the activation is still dependent on attention.

Second, the present study argues that nonnative phonemes lack entrenched articulatory templates, and therefore that the contrastive analysis of native and nonnative phonemes would allow us to verify whether early dorsal stream activity occurring during stimulus presentation is, in fact, related to automatic phonological processing. This automatic phonological processing, related to the predictive formulation of motor/articulatory mappings of likely incoming sensory signals, is likely to occur in the presence of native phonological stimuli, but not while listening to unfamiliar nonnative stimuli. In the native language, language experience has contributed to the association of sensory input with its articulatory consequences, but these processes are not likely to have been formed for unfamiliar nonnative phonemes. Greater sensorimotor activity (μ -alpha and μ -beta ERD) early during passive stimulus presentation was, therefore, hypothesized to occur during the presentation of native phonemes in comparison to nonnative phonemes. In fact, the current study results' show that passive perception of native phonemes elicited greater sensorimotor activity during stimuli presentation in comparison to nonnative phonemes, suggesting an automatic activation of auditory-articulatory mappings for native phonemes but not for nonnative phonemes. More specifically, results revealed significantly stronger alpha and especially beta ERD before native stimuli onset as well as stronger alpha ERD during native stimuli presentation. Beta ERD especially as been shown to correlate with the generation of hypothesis in relation to the phonological stimuli in order to help constrain the analysis and functionally improve speech discrimination (D. Callan et al., 2010). The fact that native and nonnative speech perception is, according to present results, differentiated by early sensorimotor activity and not by later activity (after stimuli presentation) is a strong indication that it is not being fundamentally driven by decision-making and working memory buffer related processes. Instead, our results seem to be better explained under the context of constructivist, internal model proposals of speech perception, by which this early activity is consistent with the generation of an early internal forward model, predicting motor/articulatory consequences of the incoming sensory signal. Alpha and beta ERD before and shortly during the presentation of native stimuli are consistent with the online generation of these internal models. On the other hand, stronger alpha and beta ERS before and partially during stimuli onset in the passive nonnative task are indicative of inhibitory processes (W. Klimesch et al., 1996; Pfurtscheller & Lopes da Silva, 1999), possibly resulting from gating/filtering competitive native phonemes, since no specific motor mappings exist for the nonnative stimuli. This type of activity, present during the passive task, where participants were attending to the stimuli, but were not asked to discriminate the sounds, is especially important because it enhances the possibility of automatic speech processing related to dorsal stream activity.

Our results seem to be somewhat contrary to findings by Wilson and Iacoboni (2006) in a study examining neural responses to nonnative phonemes varying in

producibility, using fMRI. In their study, they reported that nonnative phonemes activated motor brain areas more than native phonemes, in a passive task compared to rest. They interpreted the greater motor activity when hearing a nonnative phoneme to result from the continuous engagement of the motor system repeatedly attempting to access an internal auditory template that could not be readily accessed, since no match was available in the native phonological system. The incongruence of findings could be due to the different methods employed in their study. While activity in fMRI studies is measured by average activation across entire trials (Formaggio et al., 2008; Gonçalves et al., 2006; Laufs et al., 2003; Yang et al., 2010), changes in EEG rhythmic activity are quantified in time and space relative to a specific event (Jaime A. Pineda, 2005), i.e., while fMRI reported activity represents the average magnitude of the blood oxygen flow across entire trials, EEG oscillatory activity represents increases and decreases in the electrical signal in relation to a certain point in time when an event of interest is happening. Therefore, it is difficult to interpret and compare the different levels of activation in both experiments.

In active tasks, however, where attentional and decision-making resources become automatically involved, no significant differences between native and nonnative perception were found in the current experiment. When a discrimination task was involved, the attentional and decision-making processes necessary to complete the task seem to have evened out possible differences in the discrimination of native and nonnative sounds.

In summary, the current study resulted in an insightful analysis of the time-course of activation of the dorsal auditory stream while contrasting active/passive conditions with native/nonnative ones. The methodology employed provided a new understanding of the nature of the activity emanating from the dorsal stream during speech perception and its function in working memory and automatic phonological processing. The results suggest that neural processes within the dorsal auditory stream are functionally and automatically involved in speech perception mechanisms. While its early activity (during stimulus presentation) seems to be importantly involved with the instantiation of predictive motor/articulatory internal models that help constraining speech discrimination, its later activity (post-stimulus presentation) may be essential in the maintenance of working memory processes.

Limitations and Future Directions

The native and nonnative stimuli employed in this study present a voicing contrast: voiced vs. voiceless labiodental fricatives in the native condition and voiced vs. voiceless velar fricatives in the nonnative condition. A study by Murakami et al. (2015), however, suggests that sensorimotor mapping via the dorsal stream may be especially tuned to the place of articulation of perceived consonants, whereas processing underlying voicing analysis may be more related to the bilateral ventral stream. Although our study's results show sensorimotor activity within the dorsal stream related to both the discrimination of native and nonnative voicing contrasts, Murakami's (2015) study

suggests that using a place of articulation distinction could possibly have shown stronger effects.

In the present study, μ activity is interpreted as emanating mainly from the PMC (mean source location of the μ clusters was identified within Brodmann area 6), but it isn't entirely clear that sources of activation are restricted to that area since we included areas of activity emanating not just from the PMC, but also the primary motor and the somatosensory cortices. Moreover, although our mean residual variance for dipole localization was considerably low, a generic spherical head model was used, limiting the source localization. Therefore, using methods that have better spatial localization would be valuable – MEG or EEG combined with fMRI.

Also, this study could only count with the participation of female subjects. It would be useful to extend the analysis to male subjects as well.

In the future, it would be extremely interesting to further explore the inhibitory activity observed during the nonnative conditions. In the present study, we have interpreted this activity as reflecting the gating/filtering of competitive native phonemes, but more work needs to be done to verify the validity of this interpretation. It would be useful, for example, to analyze how these inhibitory processes are reflected in BOLD activity derived from fMRI data.

Findings from the present study contribute to a body of research highlighting the multiple applications and benefits of using ICA to understand speech processing in both normal and clinical populations, given its high temporal resolution, economy and availability. In the future, it would be extremely interesting, for example, to analyze the time-course of sensorimotor activation within the dorsal stream in apraxia patients.

LIST OF REFERENCES

- Alho, J., Lin, F.-H., Sato, M., Tiitinen, H., Sams, M., & Jääskeläinen, I. P. (2014). Enhanced neural synchrony between left auditory and premotor cortex is associated with successful phonetic categorization. *Frontiers in Psychology*, 5, 394. doi:10.3389/fpsyg.2014.00394
- Alho, J., Sato, M., Sams, M., Schwartz, J. L., Tiitinen, H., & Jaaskelainen, I. P. (2012). Enhanced early-latency electromagnetic activity in the left premotor cortex is associated with successful phonetic categorization. *Neuroimage*, 60(4), 1937-1946. doi:10.1016/j.neuroimage.2012.02.011
- Arbib, M. A. (2010). Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain Lang*, 112(1), 12-24. doi:10.1016/j.bandl.2009.10.001
- Basar, E., Basar-Eroglu, C., Karakas, S., & Schurmann, M. (1999). Oscillatory brain theory: a new trend in neuroscience. *IEEE Eng Med Biol Mag*, 18(3), 56-66.
- Başar, E., Schürmann, M., Başar-Eroglu, C., & Karakaş, S. (1997). Alpha oscillations in brain functioning: an integrative theory. *International Journal of Psychophysiology*, 26(1-3), 5-29. doi:[http://dx.doi.org/10.1016/S0167-8760\(97\)00753-8](http://dx.doi.org/10.1016/S0167-8760(97)00753-8)
- Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res*, 13(1), 17-26.
- Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., & Liberman, A. M. (2001). Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain Lang*, 78(3), 364-396. doi:10.1006/brln.2001.2484
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex*, 10(5), 512-528.
- Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat Neurosci*, 7(3), 295-301. doi:10.1038/nn1198
- Boersma, P., & Weenink, D. (2011). Praat: doing phonetics by computer [Computer program] (Version 5.2.45). retrieved 28 September 2011 from <http://www.praat.org/>.

- Bowers, A., Saltuklaroglu, T., Harkrider, A., & Cuellar, M. (2013). Suppression of the μ Rhythm during Speech and Non-Speech Discrimination Revealed by Independent Component Analysis: Implications for Sensorimotor Integration in Speech Processing. *PLoS One*, *8*(8), e72024. doi:10.1371/journal.pone.0072024
- Bowers, A. L., Saltuklaroglu, T., Harkrider, A., Wilson, M., & Toner, M. A. (2014). Dynamic modulation of shared sensory and motor cortical rhythms mediates speech and non-speech discrimination performance. *Frontiers in Psychology*, *5*, 366. doi:10.3389/fpsyg.2014.00366
- Buchsbaum, B., & D'Esposito, M. (2008). Short-term and working memory systems. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (Vol. 3, pp. 237-260). Oxford: Elsevier.
- Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, *315*(5820), 1860-1862. doi:10.1126/science.1138071
- Callan, A. M., Callan, D. E., Tajima, K., & Akahane-Yamada, R. (2006). Neural processes involved with perception of non-native durational contrasts. *Neuroreport*, *17*(12), 1353-1357. doi:10.1097/01.wnr.0000224774.66904.29
- Callan, D., Callan, A., Gamez, M., Sato, M. A., & Kawato, M. (2010). Premotor cortex mediates perceptual performance. *Neuroimage*, *51*(2), 844-858. doi:10.1016/j.neuroimage.2010.02.027
- Callan, D., Callan, A., & Jones, J. A. (2014). Speech motor brain regions are differentially recruited during perception of native and foreign-accented phonemes for first and second language listeners. *Frontiers in Neuroscience*, *8*, 275. doi:10.3389/fnins.2014.00275
- Callan, D. E., Callan, A. M., Honda, K., & Masaki, S. (2000). Single-sweep EEG analysis of neural processes underlying perception and production of vowels. *Brain Res Cogn Brain Res*, *10*(1-2), 173-176.
- Callan, D. E., Jones, J. A., & Callan, A. (2014). Multisensory and modality specific processing of visual speech in different regions of the premotor cortex. *Frontiers in Psychology*, *5*, 389. doi:10.3389/fpsyg.2014.00389
- Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *Neuroimage*, *22*(3), 1182-1194. doi:10.1016/j.neuroimage.2004.03.006

- Callan, D. E., Jones, J. A., Munhall, K., Callan, A. M., Kroos, C., & Vatikiotis-Bateson, E. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport*, *14*(17), 2213-2218.
doi:10.1097/01.wnr.0000095492.38740.8f
- Callan, D. E., Kent, R. D., Guenther, F. H., & Vorperian, H. K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language & Hearing Research*, *43*(3), 721-736.
- Callan, D. E., Tajima, K., Callan, A. M., Kubo, R., Masaki, S., & Akahane-Yamada, R. (2003). Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *Neuroimage*, *19*(1), 113-124.
- Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H., & Turner, R. (2006). Song and speech: Brain regions involved with perception and covert production. *Neuroimage*, *31*(3), 1327-1342.
doi:<http://dx.doi.org/10.1016/j.neuroimage.2006.01.036>
- Catani, M., Allin, M. P., Husain, M., Pugliese, L., Mesulam, M. M., Murray, R. M., & Jones, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proc Natl Acad Sci U S A*, *104*(43), 17163-17168.
doi:10.1073/pnas.0702116104
- Chevillet, M. A., Jiang, X., Rauschecker, J. P., & Riesenhuber, M. (2013). Automatic Phoneme Category Selectivity in the Dorsal Auditory Stream. *The Journal of Neuroscience*, *33*(12), 5208-5215. doi:10.1523/jneurosci.1870-12.2013
- Chollet, F. (2000). Plasticity of the Adult Human Brain. In A. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Systems* (pp. 621-638). San Diego: Academic Press.
- Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., & Pesaran, B. (2014). Sensory-motor transformations for speech occur bilaterally. *Nature*, *507*(7490), 94-98. doi:10.1038/nature12935
<http://www.nature.com/nature/journal/v507/n7490/abs/nature12935.html#supplementary-information>
- Congedo, M., Gouy-Pailler, C., & Jutten, C. (2008). On the blind source separation of human electroencephalogram by approximate joint diagonalization of second order statistics. *Clinical Neurophysiology*, *119*(12), 2677-2686.
doi:<http://dx.doi.org/10.1016/j.clinph.2008.09.007>

- Crawcour, S., Bowers, A., Harkrider, A., & Saltuklaroglu, T. (2009). Mu wave suppression during the perception of meaningless syllables: EEG evidence of motor recruitment. *Neuropsychologia*, 47(12), 2558-2563. doi:<http://dx.doi.org/10.1016/j.neuropsychologia.2009.05.001>
- Cuellar, M., Bowers, A., Harkrider, A. W., Wilson, M., & Saltuklaroglu, T. (2012). Mu suppression as an index of sensorimotor contributions to speech processing: evidence from continuous EEG signals. *Int J Psychophysiol*, 2012 Aug;85(2):242-8. doi(2012 Apr 19), 10.1016/j.ijpsycho.2012.1004.1003.
- D'Ausilio, A., Bufalari, I., Salmas, P., & Fadiga, L. (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, 48(7), 882-887. doi:10.1016/j.cortex.2011.05.017
- Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research*, 229(1-2), 132-147. doi:<http://dx.doi.org/10.1016/j.heares.2007.01.014>
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport*, 8(4), 919-924.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*, 134(1), 9-21. doi:10.1016/j.jneumeth.2003.10.009
- Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG Sources Are Dipolar. *PLoS One*, 7(2), e30135. doi:10.1371/journal.pone.0030135
- Doppelmayr, M., Klimesch, W., Hodlmoser, K., Sauseng, P., & Gruber, W. (2005). Intelligence related upper alpha desynchronization in a semantic memory task. *Brain Res Bull*, 66(2), 171-177. doi:10.1016/j.brainresbull.2005.04.007
- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), 7126-7131. doi:10.1073/pnas.1318738111
- Formaggio, E., Storti, S. F., Avesani, M., Cerini, R., Milanese, F., Gasparini, A., . . . Manganotti, P. (2008). EEG and FMRI coregistration to investigate the cortical oscillatory activities during finger movement. *Brain Topogr*, 21(2), 100-111. doi:10.1007/s10548-008-0058-1
- Fowler, C. A. (1986). an event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*(14), 3-28.

- Galantucci, B., Fowler, C., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*(3), 361-377. doi:10.3758/bf03193857
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593.
- Gallese, V., Gernsbacher, M. A., Heyes, C., Hickok, G., & Iacoboni, M. (2011). Mirror Neuron Forum. *Perspect Psychol Sci*, *6*(4), 369-407. doi:10.1177/1745691611413392
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Boston : Houghton Mifflin.
- Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: reallocation of neural substrates. *Neuroimage*, *21*(2), 494-506. doi:10.1016/j.neuroimage.2003.09.071
- Gonçalves, S. I., de Munck, J. C., Pouwels, P. J. W., Schoonhoven, R., Kuijter, J. P. A., Maurits, N. M., . . . Lopes da Silva, F. H. (2006). Correlating the alpha rhythm to BOLD using simultaneous EEG/fMRI: Inter-subject variability. *Neuroimage*, *30*(1), 203-213. doi:<http://dx.doi.org/10.1016/j.neuroimage.2005.09.062>
- Graimann, B., & Pfurtscheller, G. (2006). Quantification and visualization of event-related changes in oscillatory brain activity in the time–frequency domain. In N. Christa & K. Wolfgang (Eds.), *Prog Brain Res* (Vol. Volume 159, pp. 79-97): Elsevier.
- Hari, R. (2006). Action-perception connection and the cortical mu rhythm. *Prog Brain Res*, *159*, 253-260. doi:10.1016/s0079-6123(06)59017-x
- Hari, R., Salmelin, R., Mäkelä, J. P., Salenius, S., & Helle, M. (1997). Magnetoencephalographic cortical rhythms. *International Journal of Psychophysiology*, *26*(1–3), 51-62. doi:[http://dx.doi.org/10.1016/S0167-8760\(97\)00755-1](http://dx.doi.org/10.1016/S0167-8760(97)00755-1)
- Hatfield, G. (2002). Psychology, Philosophy, and Cognitive Science: Reflections on the History and Philosophy of Experimental Psychology. *Mind & Language*, *17*(3), 207-232. doi:10.1111/1468-0017.00196
- Hickok, G. (2009). Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans. *J Cogn Neurosci*, *21*(7), 1229-1243. doi:10.1162/jocn.2009.21189

- Hickok, G. (2012). The cortical organization of speech processing: feedback control and predictive coding the context of a dual-stream model. *J Commun Disord*, 45(6), 393-402. doi:10.1016/j.jcomdis.2012.06.004
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron*, 69(3), 407-422. doi:<http://dx.doi.org/10.1016/j.neuron.2011.01.019>
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99. doi:10.1016/j.cognition.2003.10.011
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci*, 8(5), 393-402.
- Hickok, G., & Poeppel, D. (2015). Neural basis of speech perception. In F. B. Michael J. Aminoff & F. S. Dick (Eds.), *Handbook of Clinical Neurology* (Vol. Volume 129, pp. 149-160): Elsevier.
- Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, perception & psychophysics*, 72(5), 1218-1227. doi:10.3758/APP.72.5.1218
- Houdayer, E., Cursi, M., Nuara, A., Zanini, S., Gatti, R., Comi, G., & Leocani, L. (2016). Cortical Motor Circuits after Piano Training in Adulthood: Neurophysiologic Evidence. *PLoS One*, 11(6), e0157526. doi:10.1371/journal.pone.0157526
- Houde, J., & Nagarajan, S. (2011). Speech Production as State Feedback Control. *Frontiers in Human Neuroscience*, 5(82). doi:10.3389/fnhum.2011.00082
- Iacoboni, M. (2005). Understanding others: Imitation, language, empathy. *Perspectives on imitation: From cognitive neuroscience to social science*, 1, 77-99.
- Iacoboni, M. (2008). The role of premotor cortex in speech perception: evidence from fMRI and rTMS. *J Physiol Paris*, 102(1-3), 31-34. doi:10.1016/j.jphysparis.2008.03.003
- Jensen, O., Goel, P., Kopell, N., Pohja, M., Hari, R., & Ermentrout, B. (2005). On the human sensorimotor-cortex beta rhythm: Sources and modeling. *Neuroimage*, 26(2), 347-355. doi:<http://dx.doi.org/10.1016/j.neuroimage.2005.02.008>
- Jenson, D., Bowers, A. L., Harkrider, A. W., Thornton, D., Cuellar, M., & Saltuklaroglu, T. (2014). Temporal dynamics of sensorimotor integration in speech perception and production: independent component analysis of EEG data. *Front Psychol*, 5, 656. doi:10.3389/fpsyg.2014.00656

- Jenson, D., Harkrider, A. W., Thornton, D., Bowers, A. L., & Saltuklaroglu, T. (2015). Auditory cortical deactivation during speech production and following speech perception: an EEG investigation of the temporal dynamics of the auditory alpha rhythm. *Front Hum Neurosci*, 2015 Oct 8;9:534. doi(2015), 10.3389/fnhum.2015.00534.
- 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(2), 163-178.
- Karrasch, M., Krause, C. M., Laine, M., Lang, A. H., & Lehto, M. (1998). Event-related desynchronization and synchronization during an auditory lexical matching task. *Electroencephalogr Clin Neurophysiol*, 107(2), 112-121.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Curr Opin Neurobiol*, 9(6), 718-727.
- Kilavik, B. E., Zaepffel, M., Brovelli, A., MacKay, W. A., & Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. *Exp Neurol*, 245, 15-26. doi:10.1016/j.expneurol.2012.09.014
- Klem, G. H., Luders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl*, 52, 3-6.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, 29(2-3), 169-195. doi:[http://dx.doi.org/10.1016/S0165-0173\(98\)00056-3](http://dx.doi.org/10.1016/S0165-0173(98)00056-3)
- Klimesch, W. (2012). alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci*, 16(12), 606-617. doi:10.1016/j.tics.2012.10.007
- Klimesch, W., Doppelmayr, M., & Hanslmayr, S. (2006). Upper alpha ERD and absolute power: their meaning for memory performance. *Prog Brain Res*, 159, 151-165. doi:10.1016/s0079-6123(06)59010-7
- Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neurosci Lett*, 244(2), 73-76.
- Klimesch, W., Schimke, H., Doppelmayr, M., Ripper, B., Schwaiger, J., & Pfurtscheller, G. (1996). Event-related desynchronization (ERD) and the Dm effect: does alpha desynchronization during encoding predict later recall performance? *Int J Psychophysiol*, 24(1-2), 47-60.

- Krause, C. M., Sillanmaki, L., Koivisto, M., Saarela, C., Haggqvist, A., Laine, M., & Hamalainen, H. (2000). The effects of memory load on event-related EEG desynchronization and synchronization. *Clin Neurophysiol*, *111*(11), 2071-2078.
- Krawczyk, D. C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neurosci Biobehav Rev*, *26*(6), 631-664.
- Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A. C., . . . Mazziotta, J. C. (1997). Automated labeling of the human brain: a preliminary report on the development and evaluation of a forward-transform method. *Hum Brain Mapp*, *5*(4), 238-242. doi:10.1002/(SICI)1097-0193(1997)5:4<238::AID-HBM6>3.0.CO;2-4.10.1002/(sici)1097-0193(1997)5:4<238::aid-hbm6>3.0.co;2-4
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., . . . Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp*, *10*(3), 120-131.
- Laufs, H., Kleinschmidt, A., Beyerle, A., Eger, E., Salek-Haddadi, A., Preibisch, C., & Krakow, K. (2003). EEG-correlated fMRI of human alpha activity. *Neuroimage*, *19*(4), 1463-1476.
- Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput*, *11*(2), 417-441.
- Lee, Y. S., Turkeltaub, P., Granger, R., & Raizada, R. D. (2012). Categorical speech processing in Broca's area: an fMRI study using multivariate pattern-based analysis. *J Neurosci*, *32*(11), 3942-3948. doi:10.1523/jneurosci.3814-11.2012
- Leech, R., Holt, L. L., Devlin, J. T., & Dick, F. (2009). Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *J Neurosci*, *29*(16), 5234-5239. doi:10.1523/jneurosci.5758-08.2009
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol Rev*, *74*(6), 431-461.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1-36. doi:[http://dx.doi.org/10.1016/0010-0277\(85\)90021-6](http://dx.doi.org/10.1016/0010-0277(85)90021-6)
- Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends Cogn Sci*, *4*(5), 187-196.

- Liebenthal, E., Sabri, M., Beardsley, S. A., Mangalathu-Arumana, J., & Desai, A. (2013). Neural dynamics of phonological processing in the dorsal auditory stream. *J Neurosci*, *33*(39), 15414-15424. doi:10.1523/jneurosci.1511-13.2013
- Liu, C., & Jin, S.-H. (2015). Auditory detection of non-speech and speech stimuli in noise: Effects of listeners' native language background. *The Journal of the Acoustical Society of America*, *138*(5), 2782-2790. doi:doi:<http://dx.doi.org/10.1121/1.4934252>
- Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends Cogn Sci*, *13*(3), 110-114. doi:10.1016/j.tics.2008.11.008
- Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn Sci*, *8*(5), 204-210. doi:<http://dx.doi.org/10.1016/j.tics.2004.03.008>
- Makeig, S., Delorme, A., Westerfield, M., Jung, T.-P., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2004). Electroencephalographic Brain Dynamics Following Manually Responded Visual Targets. *PLoS Biol*, *2*(6), e176. doi:10.1371/journal.pbio.0020176
- Massaro, D. W., & Chen, T. H. (2008). The motor theory of speech perception revisited. *Psychonomic Bulletin & Review*, *15*(2), 453-457. doi:10.3758/pbr.15.2.453
- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Curr Biol*, *17*(19), 1692-1696. doi:10.1016/j.cub.2007.08.064
- Möttönen, R., Dutton, R., & Watkins, K. E. (2013). Auditory-motor processing of speech sounds. *Cereb Cortex*, *23*(5), 1190-1197. doi:10.1093/cercor/bhs110
- Möttönen, R., van de Ven, G. M., & Watkins, K. E. (2014). Attention Fine-Tunes Auditory–Motor Processing of Speech Sounds. *The Journal of Neuroscience*, *34*(11), 4064-4069. doi:10.1523/jneurosci.2214-13.2014
- Möttönen, R., & Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *J Neurosci*, *29*(31), 9819-9825. doi:10.1523/jneurosci.6018-08.2009
- Möttönen, R., & Watkins, K. E. (2012). Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology*, *26*(9), 1103-1118. doi:10.1080/02687038.2011.619515

- Murakami, T., Kell, C. A., Restle, J., Ugawa, Y., & Ziemann, U. (2015). Left dorsal speech stream components and their contribution to phonological processing. *J Neurosci*, *35*(4), 1411-1422. doi:10.1523/jneurosci.0246-14.2015
- Muthukumaraswamy, S. D., & Johnson, B. W. (2004a). Changes in rolandic mu rhythm during observation of a precision grip. *Psychophysiology*, *41*(1), 152-156. doi:10.1046/j.1469-8986.2003.00129.x
- Muthukumaraswamy, S. D., & Johnson, B. W. (2004b). Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG. *Clin Neurophysiol*, *115*(8), 1760-1766. doi:10.1016/j.clinph.2004.03.004
- Muthukumaraswamy, S. D., Johnson, B. W., Gaetz, W. C., & Cheyne, D. O. (2006). Neural processing of observed oro-facial movements reflects multiple action encoding strategies in the human brain. *Brain Res*, *1071*(1), 105-112. doi:10.1016/j.brainres.2005.11.053
- Muthukumaraswamy, S. D., Johnson, B. W., & McNair, N. A. (2004). Mu rhythm modulation during observation of an object-directed grasp. *Brain Res Cogn Brain Res*, *19*(2), 195-201. doi:10.1016/j.cogbrainres.2003.12.001
- Myers, E. (2014). Emergence of category-level sensitivities in non-native speech sound learning. *Frontiers in Neuroscience*, *8*, 238. doi:10.3389/fnins.2014.00238
- Myers, E. B., Blumstein, S. E., Walsh, E., & Eliassen, J. (2009). Inferior frontal regions underlie the perception of phonetic category invariance. *Psychol Sci*, *20*(7), 895-903. doi:10.1111/j.1467-9280.2009.02380.x
- Myers, E. B., & Mesite, L. M. (2014). Neural Systems Underlying Perceptual Adjustment to Non-Standard Speech Tokens. *J Mem Lang*, *76*, 80-93. doi:10.1016/j.jml.2014.06.007
- Myers, E. B., & Swan, K. (2012). Effects of category learning on neural sensitivity to non-native phonetic categories. *J Cogn Neurosci*, *24*(8), 1695-1708. doi:10.1162/jocn_a_00243
- 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), 1-21. doi:10.1111/1469-8986.3810001
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., . . . Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*(6615), 432-434.

- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol*, *118*(12), 2544-2590. doi:10.1016/j.clinph.2007.04.026
- Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's Region: From Action to Language. *Physiology*, *20*(1), 60.
- Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Brain Res Cogn Brain Res*, *24*(2), 190-198. doi:10.1016/j.cogbrainres.2005.01.014
- Oberman, L. M., McCleery, J. P., Hubbard, E. M., Bernier, R., Wiersema, J. R., Raymaekers, R., & Pineda, J. A. (2013). Developmental changes in mu suppression to observed and executed actions in autism spectrum disorders. *Soc Cogn Affect Neurosci*, *8*(3), 300-304. doi:10.1093/scan/nsr097
- Oberman, L. M., Pineda, J. A., & Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Soc Cogn Affect Neurosci*, *2*(1), 62-66. doi:10.1093/scan/nsl022
- Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: the mirror neuron hypothesis. *Neuropsychologia*, *46*(5), 1558-1565. doi:10.1016/j.neuropsychologia.2008.01.010
- Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex*, *22*(11), 2466-2477. doi:10.1093/cercor/bhr325
- Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., . . . Hickok, G. (2010). Hierarchical Organization of Human Auditory Cortex: Evidence from Acoustic Invariance in the Response to Intelligible Speech. *Cerebral Cortex (New York, NY)*, *20*(10), 2486-2495. doi:10.1093/cercor/bhp318
- Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. *Prog Brain Res*, *159*, 99-120. doi:10.1016/s0079-6123(06)59007-7
- Onton, J., Westerfield, M., Townsend, J., & Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neurosci Biobehav Rev*, *30*(6), 808-822. doi:10.1016/j.neubiorev.2006.06.007
- Oostenveld, R., & Oostendorp, T. F. (2002). Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull. *Hum Brain Mapp*, *17*(3), 179-192. doi:10.1002/hbm.10061

- Perry, A., & Bentin, S. (2009). Mirror activity in the human brain while observing hand movements: A comparison between EEG desynchronization in the θ -range and previous fMRI results. *Brain Res*, *1282*, 126-132.
doi:<http://dx.doi.org/10.1016/j.brainres.2009.05.059>
- Petsche, H., Kaplan, S., von Stein, A., & Filz, O. (1997). The possible meaning of the upper and lower alpha frequency ranges for cognitive and creative tasks. *Int J Psychophysiol*, *26*(1-3), 77-97.
- Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*(11), 1842-1857. doi:[http://dx.doi.org/10.1016/S1388-2457\(99\)00141-8](http://dx.doi.org/10.1016/S1388-2457(99)00141-8)
- Pfurtscheller, G., Stancák Jr, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band ? an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology*, *24*(1?2), 39-46.
doi:[http://dx.doi.org/10.1016/S0167-8760\(96\)00066-9](http://dx.doi.org/10.1016/S0167-8760(96)00066-9)
- Pillai, J. J., Araque, J. M., Allison, J. D., Sethuraman, S., Loring, D. W., Thiruvaiyaru, D., . . . Lavin, T. (2003). Functional MRI study of semantic and phonological language processing in bilingual subjects: preliminary findings. *Neuroimage*, *19*(3), 565-576.
- Pineda, J. A. (2005). The functional significance of mu rhythms: Translating “seeing” and “hearing” into “doing”. *Brain Research Reviews*, *50*(1), 57-68.
doi:<http://dx.doi.org/10.1016/j.brainresrev.2005.04.005>
- Pineda, J. A., Grichanik, M., Williams, V., Trieu, M., Chang, H., & Keysers, C. (2013). EEG sensorimotor correlates of translating sounds into actions. *Front Neurosci*, *7*, 203. doi:10.3389/fnins.2013.00203
- Poeppl, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *363*(1493), 1071-1086.
doi:10.1098/rstb.2007.2160
- Poldrack, R. A. (2000). Imaging Brain Plasticity: Conceptual and Methodological Issues— A Theoretical Review. *Neuroimage*, *12*(1), 1-13.
doi:<http://dx.doi.org/10.1006/nimg.2000.0596>
- Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(20), 7865-7870. doi:10.1073/pnas.0509989103

- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature neuroscience*, *12*(6), 718-724. doi:10.1038/nn.2331
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosci*, *27*, 169-192. doi:10.1146/annurev.neuro.27.070203.144230
- Rogalsky, C., Love, T., Driscoll, D., Anderson, S. W., & Hickok, G. (2011). Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. *Neurocase*, *17*(2), 178-187. doi:10.1080/13554794.2010.509318
- Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., . . . Weiller, C. (2008). Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A*, *105*(46), 18035-18040. doi:10.1073/pnas.0805234105
- Schwartz, J.-L., Basirat, A., Ménard, L., & Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, *25*(5), 336-354. doi:<http://dx.doi.org/10.1016/j.jneuroling.2009.12.004>
- Scott, S. K., & Wise, R. J. S. (2004). The functional neuroanatomy of prelexical processing in speech perception. *Cognition*, *92*(1-2), 13-45. doi:<http://dx.doi.org/10.1016/j.cognition.2002.12.002>
- Shahin, A. J., Picton, T. W., & Miller, L. M. (2009). Brain oscillations during semantic evaluation of speech. *Brain Cogn*, *70*(3), 259-266. doi:10.1016/j.bandc.2009.02.008
- Siegel, M., Donner, T. H., & Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nat Rev Neurosci*, *13*(2), 121-134.
- Simmonds, A. J., Wise, R. J., Collins, C., Redjep, O., Sharp, D. J., Iverson, P., & Leech, R. (2014). Parallel systems in the control of speech. *Hum Brain Mapp*, *35*(5), 1930-1943. doi:10.1002/hbm.22303
- 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. *Cereb Cortex*, *17*(10), 2387-2399. doi:10.1093/cercor/bhl147
- Specht, K. (2013). Mapping a lateralization gradient within the ventral stream for auditory speech perception. *Front Hum Neurosci*, *7*, 629. doi:10.3389/fnhum.2013.00629

- Specht, K. (2014). Neuronal basis of speech comprehension. *Hear Res*, 307, 121-135. doi:10.1016/j.heares.2013.09.011
- Sterman, M. B., Kaiser, D. A., & Veigel, B. (1996). Spectral analysis of event-related EEG responses during short-term memory performance. *Brain Topogr*, 9(1), 21-30. doi:10.1007/bf01191639
- Tamura, T., Gunji, A., Takeichi, H., Shigemasu, H., Inagaki, M., Kaga, M., & Kitazaki, M. (2012). Audio-vocal monitoring system revealed by mu-rhythm activity. *Front Psychol*, 3, 225. doi:10.3389/fpsyg.2012.00225
- Tremblay, K., Kraus, N., Carrell, T. D., & McGee, T. (1997). Central auditory system plasticity: Generalization to novel stimuli following listening training. *The Journal of the Acoustical Society of America*, 102(6), 3762-3773. doi:doi:<http://dx.doi.org/10.1121/1.420139>
- Turvey, M. T. (1974). Constructive theory, perceptual systems, and tacit knowledge. In W. B. W. D. S. Palermo (Ed.), *Cognition and the symbolic processes* (pp. xii, 450). Oxford, England: Lawrence Erlbaum.
- Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, 72(2), 385-396. doi:10.1016/j.neuron.2011.09.013
- Ulloa, E. R., & Pineda, J. A. (2007). Recognition of point-light biological motion: mu rhythms and mirror neuron activity. *Behav Brain Res*, 183(2), 188-194. doi:10.1016/j.bbr.2007.06.007
- Van Mier, H. (2000). Human Learning. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Systems* (pp. 605-620). San Diego: Academic Press.
- Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik* (Vol. 9): Voss.
- Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *J Cogn Neurosci*, 15(7), 1019-1027. doi:10.1162/089892903770007407
- Warren, J. E., Wise, R. J. S., & Warren, J. D. (2005). Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci*, 28(12), 636-643. doi:<http://dx.doi.org/10.1016/j.tins.2005.09.010>
- Weiller, C., Bormann, T., Saur, D., Musso, M., & Rijntjes, M. (2011). How the ventral pathway got lost: and what its recovery might mean. *Brain Lang*, 118(1-2), 29-39. doi:10.1016/j.bandl.2011.01.005

- Weisz, N., Hartmann, T., Muller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Front Psychol*, 2, 73.
doi:10.3389/fpsyg.2011.00073
- Wernicke, C. (1969). The Symptom Complex of Aphasia. In R. S. Cohen & M. W. Wartofsky (Eds.), *Proceedings of the Boston Colloquium for the Philosophy of Science 1966/1968* (pp. 34-97). Dordrecht: Springer Netherlands.
- Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage*, 33(1), 316-325.
doi:<http://dx.doi.org/10.1016/j.neuroimage.2006.05.032>
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat Neurosci*, 7(7), 701-702.
doi:10.1038/nn1263
- Yang, L., Liu, Z., & He, B. (2010). EEG-fMRI reciprocal functional neuroimaging. *Clinical Neurophysiology*, 121(8), 1240-1250.
doi:<http://dx.doi.org/10.1016/j.clinph.2010.02.153>
- Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., . . . Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *Neuroimage*, 46(1), 226-240.
doi:<http://dx.doi.org/10.1016/j.neuroimage.2009.01.028>
- Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. i. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *Neuroimage*, 26(3), 703-720.
doi:<http://dx.doi.org/10.1016/j.neuroimage.2005.02.040>
- Zhuang, P., Toro, C., Grafman, J., Manganotti, P., Leocani, L., & Hallett, M. (1997). Event-related desynchronization (ERD) in the alpha frequency during development of implicit and explicit learning. *Electroencephalogr Clin Neurophysiol*, 102(4), 374-381.

VITA

Daniela Santos-Oliveira was born in Portugal, in 1981. After graduating from High School, in 1998, she attended the University of Aveiro, Portugal, where she graduated in 2004 as a certified teacher of Portuguese, Latin and Greek. Beside her passion for teaching, during her Bachelor's studies Daniela became especially interested in Linguistics. This interest lead her to pursue a Masters' degree in Portuguese Studies, with a concentration in Linguistics. Focusing on Portuguese phonetics, Daniela graduated in 2007 from the same university, with a thesis entitled "Analysis and Modeling of Duration in European Portuguese". In 2008, Daniela was accepted into the Speech Science Ph.D. program at the University of Tennessee Health Science Center (UTHSC). In 2009, she was awarded a scholarship from the Portuguese national funding agency for science, research and technology (www.fct.pt) to support her research towards the completion of the Ph.D. program. Her research has now been focused on the field of second language acquisition, and has been presented at the ASHA Annual Convention and the UTHSC Graduate Research Day, where she won a Best Poster award. She has also published in the Clinical Linguistics and Phonetics journal. Daniela will graduate from UTHSC in May 2017.