



8-2012

# Independent Component Analysis of Event-Related Electroencephalography During Speech and Non-Speech Discrimination: : Implications for the Sensorimotor $\Delta\infty$ Rhythm in Speech Processing

Andrew Lee Bowers

*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](#), and the [Speech Pathology and Audiology Commons](#)

## Recommended Citation

Bowers, Andrew Lee , "Independent Component Analysis of Event-Related Electroencephalography During Speech and Non-Speech Discrimination: : Implications for the Sensorimotor  $\Delta\infty$  Rhythm in Speech Processing" (2012). *Theses and Dissertations (ETD)*. Paper 28. <http://dx.doi.org/10.21007/etd.cghs.2012.0032>.

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](mailto:jwelch30@uthsc.edu).

---

# Independent Component Analysis of Event-Related Electroencephalography During Speech and Non-Speech Discrimination: : Implications for the Sensorimotor $\Delta\infty$ Rhythm in Speech Processing

**Document Type**

Dissertation

**Degree Name**

Doctor of Philosophy (PhD)

**Program**

Audiology and Speech Pathology

**Research Advisor**

Tim Saltuklaroglu, Ph.D.

**Committee**

Daniela Corbetta, Ph.D. Ashley Harkrider, Ph.D. Deborah von Hapsburg, Ph.D

**DOI**

10.21007/etd.cghs.2012.0032

**INDEPENDENT COMPONENT ANALYSIS OF EVENT-RELATED  
ELECTROENCEPHALOGRAPHY DURING SPEECH AND NON-SPEECH  
DISCRIMINATION: IMPLICATIONS FOR THE SENSORIMOTOR  $\mu$  RHYTHM  
IN SPEECH PROCESSING**

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  
Andrew Lee Bowers  
August 2012

Copyright © 2012 by Andrew Lee Bowers.  
All rights reserved.

## **DEDICATION**

I would like to dedicate this this work to my family, both old and new, without whom I would not have had the support I needed to finish it. My mother and father deserve a special dedication for providing a lifetime of unfaltering love and confidence without regard for the decisions I made. I would also like to dedicate this work to my little brother who, like me, was a student of the human brain.

## ACKNOWLEDGEMENTS

I would like to tender a general thanks to all those who provided wise council when I didn't want to continue with the long process of finishing a doctoral education. I appreciate the ECU crew, without whom I wouldn't know anything about stuttering. I would like to thank my friend Matt who contributed many hours of electrode preparation, a well-run electrophysiology lab, and Ph.D. family dinners. Dr. Stephen Crawcour should be acknowledged for his inimitable personality, skepticism, and those lengthy expostulation and reply sessions. I appreciate the guidance and support of my committee members Dr. Tim Saltuklaroglu, Dr. Ashley Harkrider, Dr. Deborah von Hapsburg, and Dr. Daniela Corbetta. My advisor and friend Tim deserves a special acknowledgment for providing a strong academic template for me to follow, hosting pub celebrations for work well done, and preparing me for the work ahead. I would like to thank my wife Lisa for her strong faith in God, tolerance, and many hours of attention to my manuscript. But my greatest debt is to Lisa, because without her faith I would never have found strength in my own.

## ABSTRACT

**Background:** The functional significance of sensorimotor integration in acoustic speech processing is unclear despite more than three decades of neuroimaging research. Constructivist theories have long speculated that listeners make predictions about articulatory goals functioning to weight sensory analysis toward expected acoustic features (e.g. analysis-by-synthesis; internal models). Direct-realist accounts posit that sensorimotor integration is achieved via a direct match between incoming acoustic cues and articulatory gestures. A method capable of favoring one account over the other requires an ongoing, high-temporal resolution measure of sensorimotor cortical activity prior to and following acoustic input. Although scalp-recorded electroencephalography (EEG) provides a measure of cortical activity on a millisecond time scale, it has low-spatial resolution due to the blurring or mixing of cortical signals on the scalp surface. Recently proposed solutions to the low-spatial resolution of EEG known as blind source separation algorithms (BSS) have made the identification of distinct cortical signals possible.

The  $\mu$  rhythm of the EEG is known to briefly suppress (i.e., decrease in spectral power) over the sensorimotor cortex during the performance, imagination, and observation of biological movements, suggesting that it may provide a sensitive index of sensorimotor integration during speech processing. Neuroimaging studies have traditionally investigated speech perception in two-forced choice designs in which participants discriminate between pairs of speech and non-speech control stimuli. As such, this classical design was employed in the current dissertation work to address the following specific aims to: 1) isolate independent components with traditional EEG signatures within the dorsal sensorimotor stream network; 2) identify components with features of the sensorimotor  $\mu$  rhythm and; 3) investigate changes in time-frequency activation of the  $\mu$  rhythm relative to stimulus type, onset, and discriminability (i.e., perceptual performance). In light of constructivist predictions, it was hypothesized that the  $\mu$  rhythm would show significant suppression for syllable stimuli prior to and following stimulus onset, with significant differences between correct discrimination trials and those discriminated at chance levels.

**Methods:** The current study employed millisecond temporal resolution EEG to measure ongoing decreases and increases in spectral power (event-related spectral perturbations; ERSPs) prior to, during, and after the onset of acoustic speech and tone-sweep stimuli embedded in white-noise. Sixteen participants were asked to passively listen to or actively identify speech and tone signals in a two-force choice same/different discrimination task. To investigate the role of ERSPs in perceptual identification performance, high signal-to-noise ratios (SNRs) in which speech and tone identification was significantly better than chance (+4dB) and low SNRs in which performance was below chance (-6dB and -18dB) were compared to a baseline of passive noise. Independent component analysis (ICA) of the EEG was used to reduce artifact and source mixing due to volume conduction. Independent components were clustered using measure product methods and cortical source modeling, including spectra, scalp

distribution, equivalent current dipole estimation (ECD), and standardized low-resolution tomography (sLORETA).

**Results:** Data analysis revealed six component clusters consistent with a bilateral dorsal-stream sensorimotor network, including component clusters localized to the precentral and postcentral gyrus, cingulate cortex, supplemental motor area, and posterior temporal regions. Time-frequency analysis of the left and right lateralized  $\mu$  component clusters revealed significant ( $pFDR < .05$ ) suppression in the traditional beta frequency range (13-30Hz) prior to, during, and following stimulus onset. No significant differences from baseline were found for passive listening conditions. Tone discrimination was different from passive noise in the time period following stimulus onset only. No significant differences were found for correct relative to chance tone stimuli. For both left and right lateralized clusters, early suppression (i.e., prior to stimulus onset) compared to the passive noise baseline was found for the syllable discrimination task only. Significant differences between correct trials and trials identified at chance level were found for the time period following stimulus offset for the syllable discrimination task in left lateralized cluster.

**Conclusions:** As this is the first study to employ BSS methods to isolate components of the EEG during acoustic speech and non-speech discrimination, findings have important implications for the functional role of sensorimotor integration in speech processing. Consistent with expectations, the current study revealed component clusters associated with source models within the sensorimotor dorsal stream network. Beta suppression of the  $\mu$  component clusters in both the left and right hemispheres is consistent with activity in the precentral gyrus prior to and following acoustic input. As early suppression of the  $\mu$  was found prior the syllable discrimination task, the present findings favor internal model concepts of speech processing over mechanisms proposed by direct-realists. Significant differences between correct and chance syllable discrimination trials are also consistent with internal model concepts suggesting that sensorimotor integration is related to perceptual performance at the point in time when initial articulatory hypotheses are compared with acoustic input. The relatively inexpensive, non-invasive EEG methodology used in this study may have translational value in the future as a brain computer interface (BCI) approach. As deficits in sensorimotor integration are thought to underlie cognitive-communication impairments in a number of communication disorders, the development of neuromodulatory feedback approaches may provide a novel avenue for augmenting current therapeutic protocols.



## TABLE OF CONTENTS

<b>CHAPTER 1. THEORETICAL BACKGROUND AND SIGNIFICANCE.....</b>	<b>1</b>
Statement of the Problem.....	1
Neurophysiological Models of Sensorimotor Integration .....	4
The Mirror Neuron System.....	4
Dual-stream Theories.....	6
Dorsal Stream Internal Models .....	6
Theories of EEG Generation.....	9
The Arceu or $\mu$ Rhythm .....	11
Blind Source Separation Algorithms .....	15
<b>CHAPTER 2. LITERATURE REVIEW.....</b>	<b>18</b>
The Spatial and Functional Neural Correlates of Speech and Non-speech Processing.....	18
Early Neuroimaging Studies .....	18
Recent Neuroimaging Studies.....	19
Synopsis and Predictions .....	24
Aims, Questions, and Hypotheses .....	26
Aim 1: To Identify Spatially Fixed and Temporally Independent Components Associated with Speech and Non-speech Processing .....	26
Question .....	26
Hypothesis.....	26
Aim 2: To Identify Independent Components Associated with Known Features of the Sensorimotor $\mu$ Rhythm.....	27
Question .....	27
Hypothesis.....	27
Aim 3: To Investigate Differential $\mu$ Rhythm Suppression Relative to Stimulus Onset, Stimulus Type, and Discriminability.....	27
Questions.....	27
Hypotheses.....	27
<b>CHAPTER 3. METHODOLOGY.....</b>	<b>29</b>
Experimental Design.....	29
Methods .....	30
Participants.....	30
Stimuli.....	30
Procedure .....	32
Data Acquisition .....	32
Data Analysis.....	33
ICA Preprocessing .....	33
sLORETA Source Estimations .....	34
Independent Component Clustering .....	35
Event-related Spectral Perturbations.....	35

<b>CHAPTER 4. RESULTS.....</b>	<b>37</b>
Behavioral Performance .....	37
Percentage Correct .....	37
Response Time.....	37
Electrophysiological Measures .....	39
Independent Component Clustering .....	39
Event-related Spectral Perturbations: Left $\mu$ .....	44
Event-related Spectral Perturbations: Right $\mu$ .....	46
<b>CHAPTER 5. DISCUSSION .....</b>	<b>49</b>
Aim 1: To Identify Spatially Fixed and Temporally Independent Components Associated with Speech and Non-speech Processing .....	49
Aim 2: To Identify Independent Components Associated with Known Features of the Sensorimotor $\mu$ Rhythm.....	51
Aim 3: To Investigate Differential $\mu$ Rhythm Suppression Relative to Stimulus Onset, Stimulus Type, and Discriminability.....	52
Theoretical, Methodological, and Clinical Significance .....	57
Future Directions .....	60
Possible Clinical Implications.....	61
Limitations .....	62
<b>LIST OF REFERENCES.....</b>	<b>64</b>
<b>APPENDIX. SUPPLEMENTAL FIGURES .....</b>	<b>79</b>
<b>VITA.....</b>	<b>92</b>

## LIST OF TABLES

Table 1-1.	MEG studies showing time-course of activation between sensory and motor regions during speech processing. ....	13
Table 1-2.	Studies showing auditory alpha and beta ERD. ....	14
Table 2-1.	Neural correlates of speech processing with decreased spectral detail as a function of condition, region, and task. ....	21
Table 4-1.	Percentage residual variance for each ECD model, ECD x,y,z coordinates, current source density x,y,z coordinates, and component rank for the left hemisphere $\mu$ cluster. ....	42
Table 4-2.	Percentage residual variance for each ECD model, ECD x,y,z coordinates, current source density x,y,z coordinates, and component rank for the right hemisphere $\mu$ cluster. ....	43

## LIST OF FIGURES

Figure 3-1.	Sample time-line of one trial and normalized stimulus amplitude over time.....	31
Figure 4-1.	Means and standard errors for percentage correct trials and response time.....	38
Figure 4-2.	Cluster results for the left-hemisphere $\mu$ component.....	40
Figure 4-3.	Cluster results for the right-hemisphere $\mu$ component.....	41
Figure 4-4.	Mean left and right hemisphere $\mu$ time-frequency ERSPs (event-related spectral perturbations).....	45
Figure 4-5.	Measures for the left-hemisphere $\mu$ cluster.....	47
Figure 4-6.	Measures for the right-hemisphere $\mu$ cluster. ....	48
Figure A-1.	Left $\mu$ $p$ FDR in the speech and tone conditions for which an active discrimination was required.....	79
Figure A-2.	Right $\mu$ $p$ FDR values in the speech and tone conditions for which active discrimination was required.....	80
Figure A-3.	Active ERSPS for the left hemisphere $\mu$ cluster.....	81
Figure A-4.	Proposed model of sensorimotor interactions in a two-forced choice discrimination task.....	82
Figure A-5.	Cluster results for the left-hemisphere $\alpha$ component.....	83
Figure A-6.	Cluster results for the right-hemisphere $\alpha$ component.....	84
Figure A-7.	Cluster results for the frontal midline $\theta$ component.....	85
Figure A-8.	Cluster results for the central midline $\theta$ component.....	86
Figure A-9.	Independent component clusters labeled according to spectral signatures and depicted on a van Essen average template.....	87
Figure A-10.	Left and right $\mu$ for correct trials in the Actsp+4dB condition averaged across 15-20Hz and 15-25Hz bands respectively.....	88
Figure A-11.	Example eye-blink components with a spatial distribution near periocular channels.....	89

Figure A-12. Example noise components with a spatial distribution near temporal channels and spectral features consistent with temporal muscle movements.....	90
Figure A-13. Example vertical eye-movement components with a spatial distribution near periocular channel. ....	91

## LIST OF ABBREVIATIONS

$\mu$	Mu
$A$	Alpha
$\theta$	Theta
BA	Brodmann's Area
BSS	Blind Source Separation
CSD	Current Source Density
dB	Decibel
DFT	Dynamic Field Theory
DLPFC	Dorsolateral Prefrontal Cortex
DRT	Direct Realist Theory
DTS	Dynamic Theory of Speech Processing
EC	Embodied Cognition
ECD	Equivalent Current Dipole
EEG	Electroencephalography
ERD	Event-related Desynchronization
ERP	Event-related Potential
ERS	Event-related Synchronization
ERSPS	Event-related Spectral Perturbations
FFT	Fast Fourier Transform
fMRI	Functional Magnetic Resonance Imaging
Hz	Hertz
ICA	Independent Component Analysis
IFG	Inferior Frontal Gyrus
MEG	Magnetoencephalography
MNS	Mirror Neuron System
MS	Motor System
M1	Primary Motor Cortex
PCA	Principal Component Analysis
PET	Positron Emission Tomography

PMC	Premotor Cortex
RT	Response Time
sLORETA	Standardized Low-resolution Electromagnetic Tomographic Analysis
STG	Superior Temporal Gyrus
STS	Superior Temporal Sulcus
vPMC	Ventral Premotor Cortex

## CHAPTER 1. THEORETICAL BACKGROUND AND SIGNIFICANCE

### Statement of the Problem

It is well known that the acoustic speech signal does not directly map onto perceived speech sound categories. In other words, many acoustic cues may be associated with the same phoneme, suggesting that acoustic cues alone cannot specify what humans perceive as distinct phonemic units. For example, although listeners identify the consonant /d/ as the same phonemic category for the syllables /di/ and /du/, the formant transitions characterizing the consonants are different for the two syllables. Despite this variability in the acoustic signal, humans successfully process speech signals in the presence of high levels of ambient noise and considerable variability in speech production and its acoustic consequences across speakers. These feats are accomplished despite the lack of a one-to-one relationship between acoustic features of the speech signal and how it is perceived by listeners. As such, it is as yet unclear how humans and other species are capable of recovering invariant (i.e., constant) information from these highly variable acoustic cues. This problem is not unique to speech perception and is referred to as the lack of invariance or perceptual constancy problem (Handel, 1989).

The primary goal of speech perception theory and more than four decades of research has been to account for the problem of perceptual invariance (Kent, 1997; Liberman, 1957; Liberman & Mattingly, 1985; Liberman & Whalen, 2000). Over the past three decades, debate about how invariant information is extracted from the acoustic speech signal has shifted from an emphasis on psychoacoustic experimental phenomena to the neural and computational basis of speech processing. As acoustic cues do not map directly to speech sound categories, neurophysiological frameworks have proposed that integrated multisensory and articulatory systems play a critical role in aiding acoustic analysis of speech. These models have been heavily influenced by the contrasting philosophical approaches of constructivism and direct-realism (Callan, Callan, Gamez, Sato, & Kawato, 2010; D'Ausilio, Bufalari, Salmas, Busan, & Fadiga, 2011; Helmholtz, 1867; Kent, 1997).

Constructivist approaches to perception suggest that the analysis of real-time sensory input is constrained by a process of internal simulation. According to this theoretical approach, previous experiential knowledge derived from speech production (i.e., the motor system) and other multisensory information (e.g., visual mouth movements) may disambiguate cues derived from the acoustic speech signal (Callan et al., 2010; Dick, Goldin-Meadow, Hasson, Skipper, & Small, 2009; Hasson, Skipper, Nusbaum, & Small, 2007; Helmholtz, 1867; Skipper, Goldin-Meadow, Nusbaum, & Small, 2007; Skipper, Nusbaum, & Small, 2005; Skipper, Nusbaum, & Small, 2006; Skipper, van Wassenhove, Nusbaum, & Small, 2007). These accounts of speech processing are similar to models of visual perception in which active hypotheses are tested against the inherently ambiguous information available to the retina (Hatfield, 2002). As such, these theories do not propose a direct match between the acoustic signal and perception. Rather, it is suggested that accurate perception is achieved via a process



in which hypothesized interpretations of the incoming speech signal weight sensory perception toward expected acoustic features.

By contrast, the direct realist approach to perception suggests that properties in sensory stimulation that have parity with self-generated actions (i.e., actions within the individual's repertoire) may be used to afford perceptual recognition (Gibson, 1986). Put another way, although the information structure from hearing, seeing, and touching is analyzed via sensory systems, what is perceived through those senses is not the structure of the information itself, but the environmental event that caused the structure. As such, the goal of all perceptual analyses is to perceive the distal, causal event. In regards to speech perception specifically, because the causal structures (i.e., phonetic gestures) are complexly encoded in the acoustic signal (i.e., acoustic correlates), the goal of perception is to directly relate the acoustic structure to the phonetic events that gave rise to them. Therefore, the direct realist approach implies that, although invariant cues are not present in the acoustic signal, sensory properties directly specify invariant articulatory gestures (Fowler, n.d.; Fowler, 1996; Fowler, Brown, Sabadini, & Weihing, 2003). As acoustic information is directly related to the motor commands that make up phonetic units, no inferential or predictive coding is required.

Due to the predictions about the computational process by which invariance is achieved, constructivist and direct-realist perspectives make clear, divergent predictions about when sensorimotor integration should occur relative to the onset of acoustic stimuli. Constructivists predict that motor and corresponding sensory activity should occur prior to and following the arrival of acoustic speech signals (i.e., hypothesis and test). Direct realists predict sensory and motor activation during acoustic stimulation only (i.e., direct matching). Although both approaches make divergent predictions, to date there has been little evidence clearly favoring constructivist over direct-realist accounts of sensorimotor integration (Callan et al., 2010). A resolution to this fundamental question may favor one computational mechanism over another for solving a vexing, decades old problem. Moreover, such a resolution may be critical to developing computational models that can accurately account for the invariance problem in dynamic, real-world contexts. Further, if sensorimotor integration does function as constructivists predict, early-cortical signals may be used to develop neurofeedback protocols designed to enhance subsequent perceptual performance in populations with communication disorders (Callan et al., 2010; Callan et al., 2003). Deficits in sensorimotor integration are thought to underlie cognitive-communication impairments in autism (Oberman, Pineda, & Ramachandran, 2007), stuttering (Kalinowski & Saltuklaroglu, 2003), hearing impairment (Le Bel, Pineda, & Sharma, 2009), and various underlying processes in aphasia (Fridriksson et al., 2009; Hickok, Houde, & Rong, 2011).

Several factors in relating theory to methodology may account for the lack of evidence favoring constructivist or direct-realist approaches. First, the past two decades of research on the neural underpinnings of receptive speech processing have focused on methods with high spatial but somewhat limited temporal resolution such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). Although fMRI and PET approaches have provided a wealth of information about the sensory and

motor cortical regions involved in speech processing, they have provided only speculation about function. Second, electrophysiological approaches such as scalp-recorded electroencephalography (EEG) capable of higher temporal-resolution have focused on event-related potentials, a measure that is time and phase-locked to the onset of a given stimulus. As this traditional measure generally provides only information about brain activity following a given stimulus event, divergent theoretical predictions about function given by constructivist and direct-realist models have not been addressed. A third factor placing limits on current electrophysiological methodology is low spatial resolution. Due to the blurring or mixing of distinct physiological sources measured via scalp-recorded EEG, activity related to cortical regions known to be involved in speech processing cannot easily be distinguished. Fourth, as few studies have demonstrated that measures of sensorimotor activity are related to perceptual accuracy, it is unclear whether sensorimotor integration performs a functional role in the perception of speech.

Recently proposed solutions to the difficulties with EEG interpretation, including blind source separation algorithms (BSS) and millisecond resolution time-frequency analysis, have made it possible to measure regional cortical activation surrounding a stimulus event (Onton & Makeig, 2006). Neuroimaging studies have traditionally investigated speech perception in two-forced choice designs in which participants discriminate between pairs of speech and non-speech control stimuli. As such, this classical design was employed in the current dissertation work using BSS and time-frequency analysis of scalp-recorded EEG to address the following specific aims to: 1) isolate independent components with traditional EEG signatures related to speech and non-speech acoustic processing; 2) identify components with features of the sensorimotor  $\mu$  rhythm, a known EEG correlate of sensorimotor integration; and 3) investigate changes in time-frequency activation of the  $\mu$  rhythm relative to stimulus type, onset, and discriminability (i.e., perceptual performance).

The primary aim of the following review is to develop a general theoretical framework within which to predict and interpret the spatiotemporal activation patterns associated with sensorimotor integration in speech processing. As no such framework currently exists, the review will attempt to synthesize predictions from speech perception theory, neurophysiological models of speech processing, and theories of EEG generation. First, relevant neurophysiological models specifying neuronal regions and computational processes involved in constructivist and direct-realist conceptual frameworks will be discussed. Second, spatial and functional neurophysiological models will be discussed with respect to current theories intended to predict and explain the neural generation of traditional EEG signatures. It will be suggested that these traditional EEG signatures, their respective sources, and their activation time-courses may be isolated using a signal separation technique known as blind source separation (BSS).

## Neurophysiological Models of Sensorimotor Integration

### The Mirror Neuron System

The current challenge for neurophysiological models of speech processing is to account for how sensory (e.g., visual and auditory) and motor systems are integrated in the central nervous system to aid in the perception of complex acoustic signals. Neuroimaging findings in humans suggest that the primary motor cortex (M1), premotor cortex (PMC), supramarginal gyrus (SMG) and the superior temporal gyrus (STG) are all active when participants listen to or watch the actions that they can also perform (see Cattaneo & Rizzolatti, 2009a). As such, it has been suggested that a neural stream now known as the human mirror neuron system (MNS) might provide a solid neurophysiological basis for sensorimotor integration (Galantucci, Fowler, & Turvey, 2006; Gallese, Gernsbacher, Heyes, Hickok, & Iacoboni, 2011). This conjecture is supported by the relatively recent discovery of a class of neurons in the F5 area of the macaque brain now known as mirror neurons. Single neuron recordings in macaques have shown that mirror neurons are active both when a primate subject observes an action and when he performs the same action, suggesting that sensorimotor integration via mirror neurons may provide a physiological basis for action understanding (see Gallese et al., 2011 for discussion).

The F5 area of the macaque is thought to be the primate homologue of the premotor cortex (PMC) near Broca's area in humans. As the PMC/Broca's is classically involved in speech production, it has been speculated neurons in this area might perform an important function in speech processing. According to these proposals, a multimodal system involved in action understanding might account for how speakers and listeners receiving multimodal sensory input translate such input into a common communicative code (i.e., parity) (Gallese et al., 2011; Rizzolatti & Arbib, 1998). In other terms, acoustic ambiguity might be resolved by linking multiple types of sensory information from speakers with the motor commands that generated the signals. Whether the MNS provides a direct link for sensorimotor integration or functions in a more dynamic manner remains controversial (Gallese et al., 2011).

Theoretical approaches favoring either constructivist or direct-realist philosophies have suggested that the MNS might provide a physiological basis for solving the problem of invariance (Gallantucci et al., 2006). As the MNS is thought to support direct links between incoming sensory information and the causal structure of that information (i.e., actions), it has been proposed that this system of neuronal regions might support direct realist accounts (Callan et al., 2010; Gallantucci et al., 2006). However, it has also been suggested that the MNS may have a more dynamic function in perception. One prominent, yet controversial notion is that the MNS not only performs associations between input modalities but also performs predictions for upcoming information, thereby bridging gaps between intended goals and missing information about those goals. For instance, primate studies have demonstrated that mirror neurons fire for completed actions (e.g., grasping a peanut) and for similar incomplete actions with the same

intended goal (e.g., moving the hand toward a peanut), suggesting that the system may code for intent or action understanding as opposed to action selection (Aziz-Zadeh & Ivry, 2009; Buccino, Binkofski, & Riggio, 2004; Buccino, Solodkin, & Small, 2006; Caggiano, Fogassi, Rizzolatti, Thier, & Casile, 2009; Cattaneo & Rizzolatti, 2009b; Corballis, 2010; Iacoboni & Dapretto, 2006; Rizzolatti & Arbib, 1998; but see Arbib, 2010 and Gallese et al., 2011 for alternative accounts). This predictive property of mirror neurons has led to speculation that the MNS may lie at the intersection of forward and inverse internal models (D'Ausilio et al., 2009a; Miall, 2003). According to these proposals, during action perception an inverse model specifies the intended communicative goal via multisensory hypotheses generated in parietal and temporal regions. Forward models are then generated in motor regions specifying the action sequence that would be required to achieve the communicative goal. In this way, inverse and forward models allow observers and listeners to simulate communicative goals within their own motor system, a process known as motor resonance. In accordance with constructivist accounts, these proposals suggest that neurons in the premotor cortex (PMC) perform important functional computations for relating multisensory information derived from a speaker with motoric information specifying communicative goals. Thus, the MNS may function not only to map speech sounds to their corresponding movements but to offer enriched comprehension by specifying communicative intent (i.e., meaning).

Recently, Arbib (2010) suggested that the MNS might accomplish this synthesis between action and meaning via a dorsal and ventral stream network. According to his computational FARS model of MNS function, a dorsal stream network maps sound onto articulatory based actions. This mechanism is thought to provide maintenance of 'parity' between auditory and motor representations of speech sounds. The network includes frontal motor areas (PMC/M1) and a region in the posterior sylvian fissure at the parietotemporal boundary (Spt). A ventral stream network supported by prefrontal and inferior temporal regions maps sound-based representations onto meaning and is connected to a distributed network supporting schemes for experiential knowledge. In accordance with constructivist theories, the two streams may cooperate to utilize experiential knowledge widely distributed throughout the brain to constrain and enrich language comprehension.

Although the participation of the MNS in sensorimotor integration offers an attractive framework through which to view computational processes involved in speech perception, the precise functional role of the MNS is controversial. Because the model predicts the mirror neurons in the PMC are critical for action understanding, it has been predicted that this region is critical for specifying phonetic units produced by a speaker and ultimately intended meaning (Arbib, 2010; Rizzolatti & Arbib, 1998). However, a number of well-known observations run contrary to this prediction, including relatively spared phonemic perception in patients with lesions restricted to the PMC/ Broca's (Hickok, Holt, & Lotto, 2009; Lotto, Hickok, & Holt, 2009) and neuroimaging findings suggesting that the PMC is not always active during speech processing tasks (Skipper et al., 2005). As such, the functional role of the MNS in speech processing is highly controversial and widely debated (Gallese et al., 2011; Lotto et al., 2009).

## **Dual-stream Theories**

Whereas MNS models of sensorimotor integration in speech processing are highly controversial, it is accepted that speech and language are processed in ventral and dorsal neural streams (Arbib, 2010; Hickok, 2009; Hickok, Houde, & Rong, 2011; Hickok & Poeppel, 2000, 2004, 2007; Rausheker & Scott, 2009). According to dual-stream models, the dorsal stream is a bilaterally organized, left-hemisphere dominant system that involves regions of the PMC, superior temporal regions (Hechyl's gyrus and STG), and a posterior temporal region for audio-motor translation (Spt). This network of regions is known as the 'how' auditory network and in receptive speech processing is thought to be involved in translating incoming acoustic information into the motor commands that might have generated phonetic units. A ventral stream, involving neural tissue in bilateral temporal regions, the dorsolateral prefrontal cortex (DLPFC) and possibly more anterior portions of Broca's area (pars triangularis), functions to map features of linguistic structure such as distinctive features, segments, phonological word forms, and semantic information (Arbib, 2010; Hickok & Poeppel, 2004; 2007; 2009; Rausheker & Scott, 2009). As such, the ventral stream is known as the 'what' network and is thought to be critical for extracting meaning from the incoming speech stream. As in Arbib (2010), the two streams are thought to interact to enrich language comprehension but the computational processes involved and the functional process by which the two both interact and compete remains a matter of speculation (Hickok & Poeppel, 2007; Skipper et al., 2006).

## **Dorsal Stream Internal Models**

Theories positing a role for sensorimotor integration in speech processing suggest that the dorsal stream network may utilize a computational process known as an internal model. Internal models were initially adapted from the motor control literature and were developed as a computational method to account for sensorimotor integration in speech production (Hickok et al., 2011). According to internal model concepts of speech production, prior to an utterance a forward model or efferent copy of the intended articulatory sequence is generated in motor regions (PMC/M1). The function of this motor copy is to inform sensory and somatosensory regions (e.g., STG and postcentral gyrus) about the likely consequences of the intended movement. An inverse model computing the desired articulator state from the forward input may then generate a corrective signal when the forward predictive model does not match with the desired articulatory state. Put another way, a projection of the desired articulatory sequence and its most likely acoustic and somatosensory outcomes are computed within the dorsal stream network. The utility of these internal models is to allow the system support speech production at natural rates. If speech production relied on slower feedback from vocalization or internal somatosensory feedback from effector state changes only (i.e., changes in muscle activity), speech could not be produced at the natural rate of between 5-10 syllables per second (Guenther, 1995, 2001, 2006). Thus, without a fast process of internal simulation via forward articulatory hypotheses it would be impossible to achieve fluent speech at natural rates (Max, Guenther, Gracco, Ghosh, & Wallace, 2004).



Internal model approaches to speech perception are consistent with constructivist philosophies and suggest that processes developed for speech production might have been ‘exapted’ at some point in human evolution for receptive speech processing. According to internal model proposals, neurons in the PMC/M1 accomplish the process of articulatory and sensory synthesis via inverse forward model pairs (IFMPs). For example, a recently proposed neurophysiological ‘motor theory’ of speech processing proposed by Skipper et al. (2006) posits that IFMPs are formed early in human development. According to that proposal, multisensory speech representations derived from sound patterns and observed facial movements are formed via the dorsal stream network. These multisensory representations are thought to constitute hypotheses about incoming sensory information from a speaker. Over the course of development, as infants begin to produce phonemes, associations between these multisensory hypotheses and speech production are formed via repeated experience (Imada & Kuhl, 2006). As in speech production, when motor commands are activated they can act as hypotheses about upcoming acoustic and somatosensory consequences via forward models or efferent copies. These motor or articulatory hypotheses may function to weight a particular interpretation of sensory signals associated with previous experience.

According to Skipper et al. (2006), the primary functional role of the motor system in this model is to resolve perception when a large discrepancy between multisensory hypotheses (i.e., acoustic and visual) and sensory analysis occurs. In other words, IFMPs may constrain the possibilities for sensory perception, allowing for accurate perception even when acoustic or visual signals are ambiguous or insufficient to specify phonemic units. It is thought that dorsal stream sensorimotor integration may play an important functional role in noisy environments, during audio-visual speech perception, and may be especially important during conversational interaction in which gestural and visual information are known to play a role in shaping perception (Hasson et al., 2007). Another important feature of this ‘motor theory’ is that internal models are instantiated only when communication partners (CPs) attend to salient multisensory, contextual, or auditory features. In other terms, during receptive speech processing in real-world communicative interaction, attention or active processing may instantiate IFMPs, weighting interpretations of the incoming multisensory speech signal toward hypotheses based on directed attention to a salient feature. In this way, it is thought that sensorimotor integration may account for invariant percepts in changing environmental conditions. A number of similar so called ‘active’ models of speech perception have been proposed by others (Rauschecker & Scott, 2009; Sams et al., 1991; van Wassenhove, Grant, & Poeppel, 2007).

As in Skipper et al. (2006), a version of Hickok & Poeppel’s ‘dual-stream’ model (Hickok & Poeppel, 2004; 2007) suggests that early forward articulatory constraints might have a functional role in perception. Recently, Hickok et al. (2011) proposed that a computational state feedback control model (SFC) might account for dorsal stream sensorimotor interactions in speech perception. This SFC framework suggests that forward models generated in motor regions may modulate perception by predicting the likely sensory consequences. According to this proposal, passive speech perception should elicit activity in the motor system only after stimulus onset, perhaps as covert

rehearsal (i.e., internal speech production) of the stimulus. However, a different mechanism is proposed to account for the functional effects of the motor system on active perception. During active perception, those authors suggest that forward predictions initiated in the motor system may generate a sensory expectation similarly to the way attention is thought to apply gain or weight to expected sensory features. In simpler terms, articulatory hypotheses may function to reduce the load on the sensory system by limiting the number of candidate perceptual targets. This model predicts early activity in the motor system and simultaneous activity in the sensory system for the generation of articulatory hypotheses or expectations, thereby limiting the number of possibilities for sensory analyses. If early forward predictions are compared with online sensory analysis, activity in the lateral portion precentral gyrus (PMC; M1) would be expected to occur prior to acoustic onset. The model also predicts that activity in the precentral gyrus should peak immediately after stimulus offset, reflecting an initial articulatory hypothesis and later synthesis with sensory information (i.e., sensorimotor integration).

Taken together, researchers proposing neurophysiological frameworks reach an accord on several important factors. A similar set of neuronal regions embodied in a mirror system or dual-stream sensorimotor network share information either important for enriching speech signals or critical to the resolving them. These regions of the brain are predicted to emerge from a dorsal stream network including the PMC, M1, STG, and Spt. If this system functions in a manner consistent with direct-realism, then it embodies a direct match between acoustic features of the speech signal and invariant articulatory gestures. By contrast, if the system functions as constructivists predict, then it likely makes use of internal model mechanisms that initially evolved or developed for sensorimotor integration in speech production. According to internal model approaches, active attention to salient features of the speech signal is important for instantiating early hypotheses that significantly shape subsequent perception. As such, internal models should perform a function in resolving speech signals during active processing. One area of contention among these theories is whether the motor system (PMC/M1) plays a critical role in speech perception depending on context (Skipper et al., 2006) or whether it has only a limited, modulatory effect (Hickok et al., 2011a).

It is also worth noting that a third conceptual framework known as embodied cognition (EC) is consistent with aspects of both Gibsonian realism and Helmholtz's concept of constructivism and has exerted increasing influence on the field of cognitive neuroscience. Whereas Gibson's concept of affordances (see Direct-realism above) emphasizes interaction between sensorimotor systems and objects within the environment, constructivism is associated with abstract schematic representations gained via repeated experience. In general terms, EC predicts that action and perceptual representation are closely interconnected and inextricable from the environmental context in which the two related systems developed, suggesting a more dynamic concept of perception than either traditional realism or constructivism provide. From this point of view, the concept of mental or perceptual representation requires redefinition. Rather than logical combinations of information contributed by cortical regions (e.g., 'Fuzzy logic' (Massaro, 1998) or a direct relationship between perceiver and environmental events, perception may be viewed as a dynamic process that is highly dependent upon

both the sensorimotor characteristics of the perceiver and goals relative to action upon the environment (Garbarini & Adenzato, 2004). Both the influence of experience emphasized by constructivists and the importance of environmental events emphasized by Gibson are encompassed by EC. As such, so called perceptual representation may be defined by the set of perceptual and motor schemas that, based on previous experience with the environment, are the same set required for successful action within the environment. Simply put, reality is not directly recreated via mental representation or directly matched with constructions based on previous experience. Instead, reality may be both filtered through previous sensorimotor learning and highly dependent on changing goals within the environment. Although this concept has not been formally applied to speech perception, in the domain of general perception EC predicts a changing role for sensorimotor integration based on the goals of the perceiver.

### **Theories of EEG Generation**

As speech is thought to be processed in a MNS or dorsal stream sensorimotor network, measureable changes in the resting state EEG would be expected during speech processing tasks. Both dual stream and MNS theories suggest that source generators of the EEG should be consistent with regional, task specific neural activation patterns within the dorsal stream. However, as EEG has traditionally been considered a low-spatial resolution measure, it is unclear how current neurophysiological models may be related EEG generation. The EEG literature is often discussed separately from the neuroimaging literature and is mired in the complexities of EEG signal processing. As such, the goal of the following section is first to discuss current theories of EEG generation. A second aim is to discuss theories of EEG generation proposing the emergence of a sensorimotor network consistent with an MNS or dorsal sensorimotor stream. The third aim is to discuss how current signal separation algorithms may be used to isolate components of the sensorimotor network.

Local/global accounts of cortical rhythm generation predict that intrinsic oscillations at central frequencies in the thalamus act as a sensory gate (Nunez, 2000). When incoming sensory information is processed, input from the thalamus results in multiple neuronal generators, each oscillating at near harmonics of thalamic rhythms (Basar, Basar-Eroglu, Karakas, & Schurmann, 2001a; Basar, Schurmann, Basar-Eroglu, & Karakas, 1997; Basar-Eroglu, Struber, Schurmann, Stadler, & Basar, 1996; Nunez, 2000; Pineda, 2005). For example, at rest the thalamus and cortex oscillate at ~10Hz. During sensory processing, it is thought that cortical neuronal generators at ~10Hz, ~20Hz, ~30Hz and ~40Hz emerge. In support of this notion, animal studies suggest that 7.5-12.5Hz oscillations in the thalamus result in the firing of cortical neurons (Steriade, McCormick, & Sejnowski, 1993). Scalp recorded EEG and magnetoencephalography (MEG) are thought to be generated by depolarizations during sensory input arriving from the thalamus. These depolarizations generate dipolar (i.e., negative and positive electrical charges) activity in cortical layers, resulting in far-field potentials that are measurable on the scalp surface (Pizzagalli, Oakes, & Davidson, 2003). These potentials are thought to be the linear sum of excitatory and inhibitory post-synaptic potentials (EPSPS/IPSPS). As



the neurons in cortical layers are oriented parallel to one another and perpendicular to the cortical surface, the sum of dipolar electric potentials can be measured on the scalp (Nunez, 2000). As such, these coherently oriented cortical columns are thought to be the primary generators of scalp-recorded EEG (Fisch, 1999). Higher frequency oscillations in the alpha (8-13Hz), beta (13-30Hz), and gamma (36-44Hz and above) bands have been associated with cognitive processes in a range of tasks and yet their functional roles are largely unknown. Controversies exist about the functional segregation of lower and upper components of these bands (e.g., 8-10Hz alpha and 11-13Hz alpha) and their possible relationships to one another (Osipova, Hermes, Jensen, & Rustichini, 2008).

Although controversies exist, the most recent reviews of activity in upper frequency bands integrate a broad range of task specific findings indicating that upper frequency oscillations predict task related performance in the domains of perception, movement, attention, memory, and complex cognition (Basar et al., 2001a; Crone, Boatman, Gordon, & Hao, 2001; Hari, 2006; Klimesch, Doppelmayr, Russegger, Pachinger, & Schwaiger, 1998; Krause, 2006; Pfurtscheller, Stancák, & Neuper, 1996). As such, it has been proposed that, whatever their relationship to one another, higher frequency components represent widespread entrainment of neural networks for task specific performance and sensorimotor processing (Basar, Basar-Eroglu, Karakas, & Schurmann, 2001b). In particular, suppression of spectral power has been shown to inversely correlate with fMRI signal change (i.e., blood oxygen level dependent measures; BOLD) in a regional manner. During synchrony or brain ‘idling state,’ it is thought that widespread synchrony between regions of activity are at rest, resulting in high spectral power and low brain metabolism in frontal, temporal, parietal, and occipital lobes. However, when regions of the cortex are involved in processing, desynchronizations (i.e., decreases in spectral power) are associated with regional increases in BOLD signal in the same regions (Formaggio et al., 2008; Goncalves et al., 2006; Laufs et al., 2003; Yang, Liu, & He, 2010).

A growing body of experimental evidence supports associations between spectral suppression and activated cortical networks involved in selective attention and motor preparation (see Pineda, 2005 for review). As the role of sensorimotor integration in speech and non-speech auditory processing is of interest in the current study, the primary aim of the following review is to discuss the possible function of the sensorimotor  $\mu$  rhythm, known to suppress during the imagination, observation, and performance of goal-directed actions (Hari, 2006; Pineda, 2005). The sensorimotor rhythm is associated with phase-locked somatosensory (~10Hz) and motor (~20Hz) components. It will be suggested first that alpha (~10Hz) and beta rhythms (~20Hz) are involved in wide-spread cognitive networks that may become entrained during active sensorimotor and cognitive processes. Although these wide-spread entrained oscillations emerge in the service of cognition, they may contribute temporally distinct information associated with short-range neuronal connections (Jung et al., 2000; Makeig, Bell, Jung, Ghahremani, & Sejnowski, 1997; Makeig, Debener, Onton, & Delorme, 2004; Makeig et al., 2002b; Makeig et al., 1999; Onton & Makeig, 2006). As will be discussed later in the chapter, this temporally distinct information may be obtained via a relatively recent data driven

statistical signal separation technique called independent component analysis (ICA) (Onton & Makeig, 2006).

### **The Arceu or $\mu$ Rhythm**

The arceu rhythm receives its name from an arch like shape and is composed of at least two dominant frequency components, one occurring at  $\sim 10$ Hz and another occurring at  $\sim 20$ Hz (Hari, 2006; Hari & Samelien, 1997; Pineda, 2005). Changes in oscillations in the arceu rhythm can be measured over the sensorimotor cortex and are brief, occurring in a time range from .2 to 2s. The  $\mu$  rhythm can be measured over the sensorimotor cortex in the absence of processing, suggesting that it might be a measure of ‘cortical idling state’ (Baar-Eroglu, Baar, Demiralp, & Schürmann, 1992; Pfurtscheller & Klimesch, 1990; Pfurtscheller et al., 1996). In other words, peak power centered near the two dominant frequencies has been observed when no processing is required. However, simple explanations of the  $\mu$  rhythm as ‘idling’ state have given way to more complex accounts of its function. For example, during processing, power in the  $\mu$  rhythm has been shown to decrease in some tasks and enhance in other tasks (Karrasch, Krause, Laine, Lang, & Lehto, 1998; Klimesch et al., 1998; Krause et al., 2000).

A large body of evidence has shown that strong sensorimotor event-related desynchronization (ERD) (i.e., power decrease or suppression) occurs during the observation of movement, imagination of movement, and the execution of movement (Pfurtscheller, 1999), strongly suggesting the  $\mu$  is a functional correlate of neural activity within the precentral gyrus. The size of ERD (often reported as % change from a baseline) is also associated with task complexity, further suggesting that  $\mu$  rhythms are important for task specific processing (Nunez & Silberstein, 2000; Nunez & Srinivasan, 2006; Silberstein, Danieli, & Nunez, 2003; Srinivasan, Bibi, & Nunez, 2006; Wingeier, Nunez, & Silberstein, 2000). Further support for this notion comes from studies of arm, mouth, and finger movements. These studies show that movement is associated with stronger  $\mu$  ERD over the sensorimotor cortex contralateral to the effector. For example, moving the left finger is associated with greater suppression over the right-hemisphere (Graiman & Pfurtscheller, 2006; Makeig et al., 2002b). Studies have also shown correspondence between the movement involved (e.g., arm vs. face movements) and distinct responses from the  $\mu$  rhythm. The  $\sim 10$ Hz component tends to be localized of the somatosensory cortex near the hand area and the  $\sim 20$ Hz component is known to emerge in a somatotopic manner from the precentral gyrus (Hari, 2006). In other words, whereas the  $\sim 10$ Hz rhythm has been localized to the somatosensory cortex, the  $\sim 20$ Hz activity of the  $\mu$  rhythm has been localized to the precentral gyrus corresponding with the primary motor cortex for the effector involved (i.e., mouth vs. finger). More lateral dipole clusters have been found for mouth movements and more medial clusters have been found for hand and finger movements (Hari, 2006; Nishitani & Hari, 2000; Salenius, Kajola, Thompson, Kosslyn, & Hari, 1995; Salenius, Salmelin, Neuper, Pfurtscheller, & Hari, 1996).

Along with suppression of the  $\mu$  rhythm, it has also been demonstrated that other source generators suppress at near harmonics of the  $\sim 10\text{Hz}$  rhythm. Laplacian derivations (i.e., spatially filtered EEG) allowing for examination of weighted scalp topography suggest independent alpha generators in frontal and occipital region (Elul, 1972). Whole head MEG further confirms the notion that alpha rhythms are generated by distinct neural networks associated with short-range neuronal connections (i.e., small patches of cortex) (Manshanden, De Munck, Simon, & Lopes da Silva, 2002). Table 1-1 provides a description of the spatiotemporal MEG correlates of receptive speech processing. It is also worth noting that an alpha-like generator in the superior temporal lobe has been recorded using MEG and intracranial recordings that is often not measured using scalp electrodes (Hari, 2006; Pineda, 2005). In accordance with a temporal generator, alpha ( $\sim 10\text{Hz}$ ) and beta ( $\sim 20\text{Hz}$ ) oscillations have been shown to be suppressed in a number of auditory processing and auditory memory tasks. Table 1-2 provides a description of experimental studies in which auditory stimuli have elicited ERD and ERS in the alpha  $\sim 10\text{Hz}$  and beta  $\sim 20\text{Hz}$  ranges.

Within the last two decades perspectives on the functional role of the rolandic  $\mu$  rhythm and diffuse alpha suppression have proposed that a common function for  $\mu$  and alpha ERD may be to transform incoming sensory information into action plans (Basar-Eroglu et al., 1996; Graimann & Pfurtscheller, 2006; Pfurtscheller et al., 1996). This framework prompted by the discovery of the previously mentioned sensorimotor neurons in the F5 area of the macaque. More specifically, this notion suggests that a MNS involved in transforming sensory input into action plans is responsible for  $\mu$  suppression during both movement observation and performance. According to this MNS hypothesis,  $\mu$  suppression localized in the primary motor cortex occurs downstream from the PMC, the proposed homologue of area F5 (Hari, 2006; Pineda, 2005). A number of experiments in which  $\mu$  rhythms were measured to perceived biological movements, have demonstrated that the  $\mu$  rhythm is reliably blocked in normal subjects during the observation of a wide range of movements, including hand, arm, mouth and even implied movement (e.g., point-light biological motion) (Crawcour, Bowers, Harkrider, & Saltuklaroglu, 2009; Muthukumaraswamy & Johnson, 2004; Muthukumaraswamy, Johnson, Gaetz, & Cheyne, 2006; Muthukumaraswamy, Johnson, & McNair, 2004; Ulloa & Pineda, 2007). Suppression of the  $\mu$  has also been shown occur less reliably in individuals thought to have dysfunction of deficit in the MNS, suggesting a functional role in movement processing (Oberman, et al., 2005; Oberman et al., 2007; Oberman, Ramachandran, & Pineda, 2008).

An elaboration of this MNS hypothesis based on information theory was also recently proposed (Pineda, 2005). According to this proposal, alpha-like rhythms serve a gating function such that broadband information is routed through narrow band filters resulting in widespread alpha entrainment when such regions are required to operate together. The thalamus is thought to serve as a broadband gate with intrinsic rhythms functioning to regulate incoming signals. This entrainment hypothesis provides an account of multiple alpha-like oscillators operating coherently and pulled toward central frequencies. However, at the same time, sources operate in an independent manner, until these regions are recruited in service of sensory analysis. During sensory processing,

**Table 1-1. MEG studies showing time-course of activation between sensory and motor regions during speech processing.**

<b>Study</b>	<b>Conditions</b>	<b>Active/ Passive</b>	<b>Spatiotemporal Pattern</b>
Nishitani & Hari (2002)	ERPs during observation and imagination of mouth movements	Active	Occipital, temporal, IFG, M1 in 70ms steps
Pulvermuller et al., (2003)	ERPs to auditory word stimuli while watching a distracting video	Passive	Nearly concurrent (separated by 20ms) IFG and STS ERP peaks
Pulvermuller et al., (2009)	ERPs during CVC and pseudo word listening relative to white noise	Active	Delay ~20ms for speech relative to noise between STS and IFG; greater amplitude in IFG for verbs
Salvi & Hari (2004)	ERPs during visual and auditory syllable discrimination	Active	Only visual speech activated primary motor cortex after stimulus onset

**Table 1-2. Studies showing auditory alpha and beta ERD.**

<b>Study</b>	<b>Conditions/Findings</b>	<b>Active/Passive</b>	<b>Spatial Pattern</b>
Krause et al., (1994)	10-12Hz alpha ERD during speech relative to non-speech	Passive	ERD across electrode locations during vowel processing relative to tone
Krause et al., (1997)	10-12Hz alpha ERD during memorization of speech played forward; ERS to speech played backward	Active	ERD across electrodes during speech forward relative to speech backward
Karrasch et al., (1998)	10-12 Hz alpha ERD during lexical matching task; ERS during auditory presentation	Active	ERD across electrodes as a function of task complexity ERS during acoustic encoding; ERD during matching
Krause et al., (1999)	10-12 Hz ERD during semantic category matching	Active	ERD across electrodes as a function of task complexity; ERS during acoustic encoding; ERD during category matching
Crone et al. (2001)	8-13 Hz alpha and gamma oscillations in syllable and tone discrimination	Active	Alpha ERD and gamma ERS during auditory processing over temporal electrodes
Cacace & McFarland (2003)	Alpha and beta oscillations in 'oddball' tone-sweep discrimination relative to passive pure tone	Active & Passive	Active tasks associated with left-hemisphere ERD at ~10Hz and ~20Hz at temporal and central electrodes relative to passive tasks
Kaiser et al., (2001)	Alpha oscillations in an 'oddball' syllable discrimination task	Active	Left hemisphere ERD during speech processing
Crawcour et al. (2009)	8-13Hz alpha range during continuous syllable counting in auditory and visual modalities	Active	Alpha desynchronization over the sensorimotor cortex during visual speech processing but not during auditory processing

independent alpha generators become entrained. The hypothesis predicts multiple alpha-like generators come into existence when sensory input requires coherent analysis of sensory information from multiple neuronal regions. As such, frontal, temporal, parietal, and occipital alpha-like generators would be predicted when all of these regions are required for sensory analysis. Perhaps most importantly, the hypothesis predicts that top-down anticipatory or active processes could bring the system into synchrony, resulting in simultaneous changes in multiple neuronal generators. Therefore, according to Pineda (2005) and others (e.g., Hari, 2006) the  $\mu$  rhythm and indeed a diffuse EEG system may represent a mechanism for translating ‘seeing’ and ‘hearing’ into ‘doing.’

## Blind Source Separation Algorithms

Interpretations of studies attempting to measure independent  $\mu$  and alpha ERD generators have been severely limited by the physical phenomenon of volume conduction. For the purposes of electrophysiology, volume conduction can be described as the transmission of electrical or magnetic fields from one current source through neural tissue to multiple sensors located on the scalp. This is a significant problem for the interpretation of EEG because fields 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. Therefore in order to test the theoretical predictions outlined in the MNS and local/global accounts of EEG generation, careful consideration must be given to the physical and electrodynamic properties of a conducting volume. The volume conduction problem has inspired skepticism about the measurement of spatiotemporal dynamics of EEG from its inception (Nunez, 2000). The next section describes relatively recent advances in EEG signal analysis and shifts in thought that make the investigation of temporally independent and spatially fixed source generators feasible.

A relatively recently proposed method for separating temporally independent neural oscillations is blind source separation implemented via independent component analysis (ICA) (Delorme & Makeig, 2004; Jung et al., 2000; Makeig, 1993; Makeig et al., 1997; Makeig et al., 2002a; Makeig et al., 1999; Onton & Makeig, 2006). Linear decompositions such as ICA and principal component analysis (PCA) separate multi-channel EEG data into maximally spatially and temporally independent components with projected 2-D scalp maps. Unlike PCA approaches, ICA does not require component projections to be orthogonal. The orthogonality requirement tends to result in so-called ‘checkerboard’ spatial maps that do not accurately represent physiological processes (Onton & Makeig, 2006). In other words, because PCA components may not spatially overlap, this process decomposes the signal into spatial maps that are not consistent with the dynamic and transient nature of neural oscillation. ICA as implemented via the automated *runica* algorithm (a version of the Infomax algorithm) in EEGLAB may be thought of as a post-processing rotation of spatial PCA components that often results in temporally independent dipolar-like scalp maps (Delorme & Makeig, 2004). A more technical description of ICA places it as a data analysis technique for reducing multivariate input into statistically independent components. ICA may be given by:



$$Y = WX$$

where  $X$  is the EEG with time in columns and electrodes in rows and  $W$  is the output of the ICA. The inverse of  $W$  ( $W^{-1}$ ) is the projection of the  $i$ th independent component (IC) onto the EEG channels, accounting for projection weights onto the original EEG data (Grimm & Pfurtscheller, 2006). ICA is entirely data driven and decomposes EEG into scalp-maps that are consistent with traditionally recognized EEG activity without taking into account head geometry or electrode location (Makeig et al., 2002a). That ICA accomplishes signal separation with no assumptions a priori strongly suggests that ICs represent distinct physiological activity related to information processing in the brain (Makeig et al., 2004; Onton & Makeig, 2006). In other terms, it is thought that because ICA identifies sources by finding maximally temporally distinct and independent sources of information in the EEG signal, these sources likely represent anatomically decoupled mechanisms.

In sum, decomposing multi-channel EEG data using independent component analysis addresses a number of the proposed difficulties with EEG interpretation. First, as noise sources from blink and other types of artifact are not temporally related to cortical activity, ICA is a robust method for removing artifact from EEG signals. Second, because IC's are considered maximally independent from one another, volume conduction from multiple sources cannot account for changes in time and frequency. Third, because with sufficient training data, ICA decomposes signals into spatially fixed and maximally temporally independent scalp maps, inferences about underlying processes between brain and behavior can be made with more confidence. Finally, although ICA does not solve the source localization problem directly, low variance has been observed between dipolar source reconstructions and ICA projections for both low and high-density electrode arrays, suggesting that it may act as an efficient spatial filter for spatial localization (Congedo, Gouy-Pailler, & Jutten, 2008; Makeig et al., 2004).

The aim of the preceding review was to merge speech perception theory, neurophysiological models of receptive speech processing, and theories of EEG generation to provide a general framework for predicting and interpreting spatiotemporal EEG patterns during speech processing. Internal model concepts of speech processing clearly predict early forward and inverse models play an important functional role in speech processing. As these models are thought to be embodied in a dorsal stream network including the PMC, M1, STG, and area Spt, source activation of all of these regions is predicted during auditory speech processing. As such, ICA decomposition of event-related EEG during a speech processing task would be expected to separate sources with topographic scalp maps and source reconstructions consistent with the dorsal stream network. Regarding the time-course of source activation patterns, if constructivist accounts provide accurate predictions, during active speech processing tasks rhythms known to be associated with sources in motor regions, such as the  $\mu$  rhythm, should desynchronize (i.e., suppress) prior to the onset of an acoustic stimulus and peak during the period in which early forward models are compared with sensory analysis (i.e., during sensorimotor integration). By contrast, if the direct-realist account of perception provides accurate predictions, early desynchronization of the  $\mu$  would not be expected. In contrast

with constructivist theory, peak suppression would be expected when incoming acoustic stimuli are directly related to invariant speech gestures (i.e., during stimulus presentation).

Although theoretical models provide a general framework within which to interpret EEG data, two decades of neuroimaging research have approached questions about the neurophysiological mechanisms underlying receptive speech processing. As such, the primary aim of chapter two is to discuss neuroimaging designs and approaches to the study of speech perception. As much of the research has addressed the controversial question of how sensory and motor subsystems function in resolving speech in the auditory modality, the main focus of the literature review concerns sensorimotor activation during auditory speech processing.



## CHAPTER 2. LITERATURE REVIEW

### **The Spatial and Functional Neural Correlates of Speech and Non-speech Processing**

In recent years, there has been an enormous increase in the number of publications using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) to study human cognition, emotion, perception and action. PET measures regional cerebral blood flow (cRBF), glucose metabolism, or oxygen consumption in neural tissue by detecting positrons emitted from a radioactive tracer. fMRI indirectly measures changes in the oxygenation of hemoglobin by detecting perturbations in a magnetic field and is thus referred to as a blood oxygen level dependent measure (BOLD). Both PET and fMRI provide 3-D pictures of regional brain activity during online thinking, perceiving, and acting. Whereas these two techniques offer millimeter spatial resolution both have limited temporal resolution. PET requires as much as 30 seconds of recording and BOLD signals can measure second-by-second regional changes in neuronal activity (Wager, Hernandez, Jonides, & Lindquist, 2007). Electrophysiological methods such as EEG and MEG have higher temporal resolution but limited spatial resolution and thus less precision in identifying the sources of neuronal activity. To address both the temporal limitations of neuroimaging methods and the spatial limitations of EEG/MEG, recent attempts have been made to combine the two approaches (Callan et al., 2010).

One disadvantage of any neuroimaging method is that only the regional and temporal correlates of neural activation during task performance can be determined. As such, in recent years another method referred to as transcranial magnetic stimulation (TMS) has been developed to establish causal links between brain and behavior. TMS involves the use of brief magnetic pulses delivered via a magnetic wand at the scalp surface to stimulate focal areas of cortical tissue. This stimulation disrupts neural activity resulting in a temporary lesion. Using TMS methodology, psychophysiological paradigms can provide information about causal relationships between disruptions in processing and behavioral performance (e.g., reaction time or accuracy measures). In recent years, the goal of a growing number of studies using neuroimaging, TMS, and other psychophysiological methods has been to combine data from human performance, neuropsychology, and psychophysiology to offer rich information about the functional capacities of the human brain. In the last three decades, neuroimaging methods have been brought to bear on the debate about sensorimotor involvement in speech and non-speech perception.

### **Early Neuroimaging Studies**

The notion that motor involvement in speech processing is highly context dependent was first suggested to explain early imaging findings. In early studies, passive speech perception tasks were correlated with bilateral temporal lobe activation consistent with sensory analysis in speech perception (Binder et al., 1994). However studies

employing classical speech perception paradigms requiring an active decision about two percepts (e.g., /ba/ vs /da/) found bilateral temporal activation along with prominent activity in frontal regions (e.g., Broca's area/PMC). One of the earliest studies to find activity in Broca's region measured rCBF during active discrimination of noise bursts and speech syllables. Whereas the noise bursts appeared to result in increased activity in the primary auditory cortex, acoustically matched speech syllables resulted in bilateral activation in the secondary auditory cortices. Further, discrimination of phonetic structure (i.e., active decision) was associated with activation in the opercular portion of Broca's area (BA44) and the premotor cortex, suggesting a role for articulatory recoding during speech perception (Zatorre, Evans, Meyer, & Gjedde, 1992). Another early study designed to delineate the functional role of frontal-motor activity (i.e., cRBF), found that activation of the PMC occurred during auditory word recognition but was greater when words were presented at a slower rate, and greater still for word repetition. Those authors concluded that, because the activity in the PMC is subtle and sensitive to differences in the task, the study design employed and imaging methods themselves are important considerations (Zatorre & Belin, 2001).

### **Recent Neuroimaging Studies**

Driven in part by theories of speech perception predicting motor involvement and bolstered by the discovery of the human MNS, more recent fMRI studies have continued to provide evidence of motor involvement in speech perception. At the sublexical level (i.e., syllable level), Wilson, Saygin, Sereno and Iacoboni (2004) found that passively listening to meaningless syllables and producing the same syllables was associated with overlapping regional activation (i.e., increase in BOLD signal) in primary motor and premotor regions. However, as other studies have not found significant activation in motor regions during passive processing (Zatorre, Meyer, Gjedde, & Evans, 1996), it has been suggested that tasks in which predictive coding is required may be better associated with sensorimotor integration in speech perception (Callan et al., 2010). Consistent with this notion, the active learning of new sensory and motor speech patterns has also been associated with regional cerebral increases in both sensory and motor regions (Golestani & Zatorre, 2004).

It has been demonstrated that second language learning of an /r/ vs /l/ contrast is associated with a relative increase in sensory temporal areas (STS/STG), posterior temporal areas (Spt), and in premotor regions (i.e., Broca's, primary and premotor cortex, and anterior insula) (Callan et al., 2003). Further, this network of neural motor, sensory, and orosensory regions was also correlated with better /r/ vs. /l/ identification performance in second language listeners, suggesting that this network of regions plays a functional role in learning new phonological contrasts. Callan et al. (2003) suggested that the function of the motor system may be to produce an early forward model that encodes the likely sensory consequences. This forward model is then compared with incoming sensory information in the temporal lobes. As in speech production models (Callan et al, 2000), the Spt region functions as a sensorimotor interface, generating error signals for predicted relative to actual sensory input.

Along with auditory learning and active discrimination it has also been demonstrated that activity in the premotor cortex/Broca's increases with the level of spectral detail in the acoustic signal. Binder, Liebenthal, Possing, Medler and Ward (2004) employed a design in which sublexical meaningless synthetic speech syllables (i.e., /ba/ and /da/) were embedded in various levels of Gaussian white-noise. Participants indicated whether the syllables sounded the same or sounded different using a two-choice button press response pad. Perceptual accuracy, response time, and BOLD signals were measured during task performance. As expected, very low signal-to-noise ratios (SNR's) were associated with decreases in perceptual accuracy (i.e., no better than chance) and increases in reaction time, whereas higher SNR's were associated with increases in accuracy and decreases in reaction time. BOLD signals in the anterior insula were found to positively correlate with response time, whereas bilateral activity in temporal regions was found to correlate with accuracy. Activity in the ventral PMC was negatively correlated with accuracy, suggesting that activity in the PMC might provide articulatory information aiding or constraining sensory analysis when spectral detail is decreased. In other words, because PMC activity decreased as accuracy increased, activity in that region may be interpreted as providing critical information for auditory analysis when spectral detail is decreased, strongly supporting a constructivist view of speech processing. Further, because activity in the PMC was not correlated with reaction time (thought to be a correlate of sensory decision processing), the study suggested that the PMC plays a functional role in sensory processing apart from sensory-decision mechanisms. A large number of studies have been dedicated to speech processing as a function of spectral detail in the speech signal. Most have concluded that the PMC may play a compensatory role when spectral detail is low, perhaps as articulatory recoding for later comparison (Hickok & Poeppel, 2007). As in previous studies, active tasks are more likely to elicit activity than passive tasks. Table 2-1 provides a review of recent studies investigating the neural correlates of speech-in-noise as a function of condition, region of activation, and active or passive task performance.

One possibility suggested early in the neuroimaging literature is that motor regions involved in sublexical discrimination tasks are related the selection and maintenance of stimuli in working memory. In one of the most cited papers in the literature, Hickok and Poeppel (2004) expanded on this notion by suggesting that activity in motor regions during sublexical phonological discrimination occurs because stored auditory percepts must be translated into articulatory code for comparison in working memory. This suggestion motivated work in which the cognitive demands of various active sublexical tasks were manipulated with the goal of showing that phonological working memory demands would increase activity in motor regions. Burton et al. (2000) demonstrated that cognitive load during speech discrimination could be increased by employing a task in which participants could not rely on a single phonetic cue to make a decision (e.g., /ket/ vs. /lim/). This task was compared to another phonological discrimination task in which only one phonetic cue was needed for comparison (/pet/ vs. /bet/). Tasks in which more than one cue must be used are considered segmentation tasks, whereas tasks requiring the comparison of only one cue are considered identification tasks. Consistent with the notion that the left IFG/PMC was involved in articulatory recoding, the PMC (BA44) was more active for the former task (i.e., more than one cue)

**Table 2-1. Neural correlates of speech processing with decreased spectral detail as a function of condition, region, and task.**

<b>Study</b>	<b>Conditions/Stimuli</b>	<b>Task</b>	<b>PMC</b>	<b>STG/STS</b>	<b>Spt</b>
Salvi et al., (2002)	Quiet and monoaurally presented speech in broadband noise	Active	X	X	
Scott et al., (2004)	Spoken sentences in broadband and multitalker babble	Passive		X	
Wong, et al., (2008)	Word listening and active matching to pictures in +20 and -5db SNR	Active	X	X	X
Wong et al., (2009)	Word listening and active matching to pictures in young and older groups in +20 and -5db noise	Active	X	X	X
Scott et al., (2006)	Noise vocoded speech (speech in which noise is inserted) passive sentence listening	Passive		X	
Oblesser et al., (2010)	Word listening in broad band noise	Active	X	X	X
Ones et al., (2010)	Spectral detail manipulated from broadband noise to noiseless speech	Passive	X	X	

when compared to the later task (i.e., one cue). However, if activity in BA44 is due primarily to demands required for articulatory recoding, then a similar tone perception task requiring similar demands on working memory should not elicit activity in BA44. To that end, Burton and Small (2006) developed a tone perception task in which a sequence of three tones was generated. These tone sequences were made up of pairs differing for all three different tones and the task was to indicate whether the initial tone in the sequence was the same or different. The tone segmentation task was compared to the speech segmentation task used in Burton et al. (2000). Compared to a resting baseline, both the tone and speech segmentation tasks resulted in activation bilaterally in the temporal lobes and in the left middle frontal gyrus (MFG) bordering on the IFG, suggesting that articulatory recoding might not be the only function of the motor system. Consistent with previous neuroimaging studies (Zatorre & Binder, 2000), tone discrimination was associated with increases in right-hemisphere frontal regions as well. Left hemisphere IFG/PMC activation has also been shown in active rapid pitch discrimination tasks, suggesting that it might have a function in the processing of rapidly changing auditory signals generally (Johnisse & Gati, 2003). In addition, LoCasto et al. (2004) found that words, pseudowords, and tone sequences resulted in greater activity in the IFG (BA45/46). Neither discrimination performance nor activation patterns were different for words relative to pseudowords, suggesting that lexical access did not tax processing for working memory. Thus, in addition to non-native phoneme learning, non-native contrast discrimination, and active listening in degraded speech processing tasks, motor regions appear to have functions for phonological working memory or general working memory for rapidly changing acoustic signals.

Combined time-frequency (i.e., MEG/EEG) and fMRI approaches have shown strong correlations between localized frequency modulations and spatial patterns in speech processing (Giraud et al., 2007). In a passive speech processing, Giraud (2007) showed correlations between BOLD signals and EEG. In the left auditory cortex (i.e., Heschl's), high frequency endogenous rhythms (28-40Hz) were highly correlated with BOLD signals, whereas low frequency rhythms were correlated with right hemisphere activity in the same region (3-6Hz). In addition, higher rhythms (28-40Hz) were associated with activity in the PMC whereas lower rhythms (3-6Hz) were associated with activity corresponding with the tongue region, suggesting that the auditory and motor systems are coupled. In another combined fMRI/MEG study, Callan et al. (2010) employed a classical single syllable speech perception task embedded in a two-forced choice anticipation paradigm using fMRI and subsequent constrained MEG methods. Both active and passive syllable identification tasks were compared to listening in white noise. As the speech syllables were also embedded in white noise (+4dB SNR), levels of correct syllable identification (i.e., same or different) were reduced by 20-30%. This manipulation allowed for trials in the active task to be separated into correct and incorrect trials. The first fMRI study revealed that only the premotor cortex (BA44/45) was associated with correct trials relative to baseline, passive listening, and incorrect active listening. The second MEG study revealed that early (i.e., prior to perception) and late (~200ms after perception) premotor time-frequency oscillations in the upper alpha, lower beta, and gamma range were associated with correct relative to incorrect performance and passive perception, strongly indicating that the premotor cortex plays an essential,

functional role in active speech discrimination tasks. The Callan et al. (2010) study is the first to provide strong evidence for a top-down functional role for the premotor cortex in speech perception.

In addition to functional imaging, TMS studies have approached the hypothesis that motor systems are critical for speech perception by systematically deactivating the premotor and sensory areas. To that end, these studies used repetitive TMS (rTMS) to cause virtual or temporary lesions in motor and sensory regions (Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007; Sato, Rondinoni, Sturzbecher, de Araujo, & Amaro, 2010). Meister et al. (2007) applied rTMS to the primary motor, PMC, and to the left STG before participants actively discriminated speech syllables embedded in white noise. The study found that disrupting the PMC but not the STG resulted in a decrease in perceptual performance, suggesting that the premotor cortex was essential for performance. However, because that study used speech-in-noise stimuli only, it was suggested that the premotor cortex played an essential role only when the spectral detail was reduced and percepts were ambiguous. Sato et al. (2009) demonstrated that deactivation of the same premotor region resulted in decrease in reaction time on a speech segmentation task relative to a speech identification task. Those findings were consistent with those of Locasto et al. (2006), in which increases in the demand for phonological working memory were associated with increased activation of the PMC. In the most recent study to date, (D'Ausilio et al., 2009b) found that repeated stimulation of the primary motor cortex in a somatotopic manner increased the perception of speech sounds requiring lip and tongue movement respectively, strongly suggesting that the motor cortex modifies speech perception. Taken together, TMS studies strongly indicate a functional role for the premotor and motor cortices in sublexical speech processing.

Although a number of studies have found activity in the PMC/M1 during the perception of speech, the precise functional role of sensorimotor integration is still a matter of heated debate (Gallese et al., 2011). First, imaging studies have shown considerable variability in the precise region of the motor system involved in auditory perception, suggesting that such activation may be dynamic and highly dependent on context (Price, 2010). A recent activation-likelihood meta-analysis found that a region superior and anterior to PMC was active in sublexical speech perception tasks (Turkeltaub & Coslett, 2010). Second, debate about the role of the PMC relative to sensory systems in receptive processing and whether it is part of a MNS for speech perception is ongoing. Concepts designed to account for sensorimotor integration propose either greater weight for the sensory-oro-sensory system or co-equal participation in the perception of speech (Gallese et al., 2011). At a minimum, sensorimotor integration plays a modulatory role in sensory processing in few, selective situations (Hickok et al., 2011) and more likely plays an essential role, perhaps depending on the environmental context in which processing occurs (D'Aussillo et al., 2009; Pulvermüller & Fadiga, 2010). Rauschecker and Scott (2009) and Skipper et al. (2006) emphasize that an internal simulation system likely preforms crucial functions during face-to-face interaction as listeners track a number of communicative gestures generated by the speaker, allowing parity between sender and receiver of information.



As in Callan et al. (2003) and others (e.g., Davis & Johnstrude, 2007; Hickok et al., 2011; Price, 2010; Skipper et al., 2006) it appears that one candidate function for the motor system in active auditory perception is to place early constraints on incoming sensory signals. These signals are then compared in a region thought to be involved in sensorimotor integration (i.e., area Spt). A similar hypothesis, suggesting a relationship between selective attention and the motor system has recently been proposed by Hickok et al. (2011), in which motor-selective attention facilitates processing in active tasks. However, testing these hypotheses would require establishing a time-course for sensorimotor activation pattern related to perceptual performance. For instance, it remains possible that, rather than early articulatory models, the motor system may function to recode sensory information into articulatory information to enhance working memory (i.e., covert rehearsal hypothesis; Hickok & Poeppel, 2004; 2007). Clearly, more time-sensitive methods with spatial resolving power at the hemispheric and regional levels are required to test hypotheses about internal models.

### **Synopsis and Predictions**

Despite more than two decades of study, the spatiotemporal signatures of sublexical speech and non-speech processing are unclear. The measurement of EEG signatures of speech processing in time using BSS may yield important contributions to the extant literature. The primary goal of the first and second chapters has been to synthesize predictions from speech perception theory, neuroimaging findings, and theories of EEG generation to provide a general framework for predicting and interpreting the spatiotemporal course of EEG signatures during speech and non-speech processing. Taken together, models of speech and complex acoustic processing suggest that identifying the time-course of sensorimotor integration in relation to the onset of acoustic stimuli may shed light on how those systems function.

Although it remains unclear how the upper frequency components of the EEG are related, it is clear that such oscillations are important for the processing and generation of sensory and motor information. Perspectives predicting global and local entrainment of EEG rhythms suggest that the rolandic  $\mu$  rhythm and other alpha-like rhythms should emerge during sensory and motor processing when the task requires active processing or top-down control. Further, as alpha and beta rhythms have been localized to sensory, premotor, and motor regions during the performance of various tasks, it can be predicted that speech and non-speech will elicit regional changes in both upper frequency signal components. Further, as both rhythms are thought to be important for the processing of sensory stimuli differences as a function of task complexity and performance level would be expected.

Regarding the spatiotemporal activation the dorsal stream during an anticipatory sublexical forced choice task, several predictions can be made from current views of EEG generation (Pineda, 2005). During a passive sublexical discrimination task in which top-down modulation is not required,  $\mu$  ERD (~10Hz and ~20Hz) in motor regions may occur only after sensory analysis or not at all. During active tasks, in which top-down

processing in frontal-motor regions may be important for task performance, suppression of the sensorimotor rhythm and other alpha-like rhythms should occur prior to and immediately after sensory analysis. As it is thought that EEG time-frequency components become entrained for active sensory analysis, it can be predicted that changes in  $\mu$  ERD and other alpha frequency components will be found in sensorimotor, temporal, and possibly posterior temporoparietal regions. Therefore, during active tasks, ICA is expected to reveal alpha, beta, and perhaps low gamma spectral changes (i.e., suppression) in left and right hemisphere frontal, temporal, and perhaps temporoparietal regions.

Constructivist and direct realist theories offer clear predictions about the time-course of activity in brain areas thought to be involved in sensorimotor integration. First, in passive auditory speech discrimination tasks (e.g., discrimination of /ba/ and /da/), although motor activation may occur along with activation in sensory areas, it may not be necessary for perceptual performance. In support of this perspective, neuroimaging studies suggest inconsistent activation of the motor regions during passive speech perception tasks and lesion evidence suggests that motor systems are not necessary for speech perception in good listening conditions. If as Hickok and Poeppel (2007) suggests, the motor activity in passive tasks is merely epiphenomenal covert rehearsal, then motor activity as indexed by the sensorimotor  $\mu$  rhythm should only occur after sensory analysis or perhaps not at all.

During active tasks, constructivist approaches predict that forward constraints on auditory analysis generated in motor regions may function either by enhancing neural gain for expected features or by suppressing neural gain for irrelevant features (Callan et al., 2010; Hickok et al., 2011). Perhaps this top-down modulation functions to enhance the baseline excitability or gain of neural populations relevant to the upcoming analysis. As such, active processing may be enhanced by internal simulations occurring prior to perception. Therefore, perceptual processing should be preceded by activity in motor regions. Finally, because it is thought that forward predictions are then checked against incoming acoustic input, suppression of the  $\mu$  rhythm would be expected during acoustic analysis with peak suppression occurring immediately following acoustic analysis. In other words, since early forward models are thought to be related to active task performance, peak activity would be expected when early hypotheses are compared with acoustic analysis (i.e., sensorimotor integration).

In regards to the processing of non-speech sounds, if motor and sensory phonological regions are both active, only activity after or during analysis would be expected. Because constructivist approaches invoke forward internal models that are based on previous experience, meaningless rapid-tone changes should not produce early motor or sensory constraints in active tasks. In other words, because a lifetime of previous experience for analyzing meaningless rapid-tone sequences does not exist, early motor models that function to constrain auditory analysis are not likely. The motor activity to rapid-auditory processing found in some previous studies may reflect a functionally distinct process (Burton et al., 2005; Joanisse & Gati, 2003; LoCasto et al., 2004a).



The predictions of constructivist approaches are in contrast with those of direct-realism. Direct-realism predicts that incoming acoustic signals directly specify the gestures that caused them. Although direct-realism does not directly specify neural system involvement it clearly predicts a relationship between acoustic analysis and gesture. Therefore, motor system involvement should occur concurrently with sensory analysis and no inferential or predictive coding would be required. Perhaps direct-realism is in the best position to explain previous findings of IFG and MFG (i.e., cognitive motor) activity during non-speech pitch perception (e.g., LoCasto, Krebs-Noble, Gullapalli, & Burton, 2004b). Because direct-realism suggests that even non-speech auditory signals that have been paired previously with actions might excite the motor cortex, it seems possible that concurrent sensory and motor activity after stimulus onset might occur. Further direct-realism predicts a direct relationship between non-speech sensory and motor analysis, in the case of pitch perception or rapid temporal processing, it remains possible that incoming acoustic signals can be transduced into motor commands that might generate the desired sensory consequences. For example, a rising pitch change might be simulated or approximated internally using a vocal tract model (Burton, 2009). In this example, motor processing should occur only during or after the onset of an acoustic event.

## **Aims, Questions, and Hypotheses**

### **Aim 1: To Identify Spatially Fixed and Temporally Independent Components Associated with Speech and Non-speech Processing**

#### **Question**

If the application of ICA results in left hemisphere independent dipolar scalp maps in frontal, motor, and temporal regions in the left-hemisphere, will equivalent current dipole models and current source density estimates (i.e., sLORETA solutions) show activity in regions consistent with known cortical regions involved in speech and non-speech processing?

#### **Hypothesis**

If ICA is driven by small patches of synchronized cortex related to stimulus processing, the application of ICA to raw EEG multichannel data will result in topographic maps ( $W^{-1}$ ) with associated ECD models and distributed CSD solutions consistent with processing in regions known to be active in sublexical speech processing tasks.

## **Aim 2: To Identify Independent Components Associated with Known Features of the Sensorimotor $\mu$ Rhythm**

### **Question**

Will independent component clustering reveal components consistent with known features of the sensorimotor  $\mu$  rhythm including spectral peaks at  $\sim 10\text{Hz}$  and  $\sim 20\text{Hz}$ , dipole locations and sLORETA solutions in the precentral and postcentral gyrus, and suppression in the  $\sim 10\text{Hz}$  and  $\sim 20\text{Hz}$  range?

### **Hypothesis**

Independent component clustering will reveal components consistent with known features of the sensorimotor  $\mu$  rhythm including spectral peaks at  $\sim 10\text{Hz}$  and  $\sim 20\text{Hz}$ , dipole locations and sLORETA solutions in the precentral and postcentral gyrus, and suppression in the  $\sim 20\text{Hz}$  range.

## **Aim 3: To Investigate Differential $\mu$ Rhythm Suppression Relative to Stimulus Onset, Stimulus Type, and Discriminability**

### **Questions**

- 1) Is the time-course of power suppression for  $\mu$  component clusters significantly different for passive speech and tone processing condition when compared to passive noise baseline ( $p\text{FDR} < .05$ )?
- 2) Is the time-course of peak power suppression for  $\mu$  component clusters different for active speech and tone processing ( $p\text{FDR} < .05$ )?
- 3) Is the time-course of peak suppression for  $\mu$  component clusters different for correct trials compared to trials in which percepts are identified at chance levels ( $p\text{FDR} < .05$ )?
- 4) Is the time-course of peak suppression for  $\mu$  component clusters significantly different for correct speech and tone perception trials ( $p\text{FDR} < .05$ ).

### **Hypotheses**

- 1) If passive speech perception is associated with sensorimotor activity,  $\mu$  suppression prior to and following stimulus onset will be significantly different from a passive noise baseline.
- 2) If constructivist, internal model perspectives provide an accurate account, early  $\mu$  suppression prior to stimulus onset and following stimulus onset should be significantly different from the passive noise baseline for the active syllable discrimination task.

- 3) If  $\mu$  suppression is related to perceptual performance, then active syllable discrimination trials in which discrimination was performed accurately should be different from trials in which discrimination was at chance levels.
- 4) If the function of motor activity is to perform top-down constraints on speech processing as opposed to a general function related to auditory processing (i.e., general attention/working memory), the time-course of  $\mu$  suppression is expected to be different for correct speech relative to correct tone perception trials.

## CHAPTER 3. METHODOLOGY

### Experimental Design

The current proposal employs a classical forced-choice speech and non-speech (i.e., rapid pitch change) perception design in which EEG time-frequency changes (i.e., ERS/ERD) can be measured before, during, and after stimulus onset (Binder et al., 2004; Callan et al., 2010; Johannisse & Gatti, 2003; Zatorre et al., 1992). As activity in the premotor and primary motor regions has been observed when spectral detail of the speech signal is reduced, speech and non-speech percepts will be embedded in Gaussian white-noise. The presence of white-noise has also been shown to decrease perceptual performance for the proposed task (Binder et al., 2004; Callan et al., 2010), allowing for an examination of correct trials and those not identified above chance levels. Such an examination is crucial to establish whether the EEG signatures of sensory and motor processing are functionally related to stimulus processing. As any auditory stimuli may result in broad changes in the EEG spectrum (e.g., Krause, 2006), speech and tone processing tasks will be compared to a baseline condition in which participants passively listen to white-noise.

Other considerations in the design are motivated by predictions about the role of motor and sensory systems as a function of task performance and concerns that have traditionally limited previous neuroimaging studies. First, as motor processing in passive tasks is thought to reflect covert rehearsal (i.e., after stimulus onset) and motor processing in active tasks is thought to play a functional role (prior to and after stimulus onset), both passive and active tasks will be employed. Second, in two forced-choice paradigms in which a phonetic decision is required, it has been suggested that activity recorded within classical speech production areas might arise from premotor planning for a button press response or working memory for sensory-decision. To address this possibility, two active tasks (i.e., requiring a sensory-decision via button press) were compared for both speech and tone stimuli, one in which good perceptual performance was expected and another in which performance was not expected to be above chance levels. If suppression of the sensorimotor rhythm were related to a button press response as opposed to stimulus processing, no differences between correct trials and those discriminated at chance levels would be expected. Further, no differences would be expected for speech and tone stimuli since for both an active decision was required.

Third, in previous neuroimaging studies, due to the limited temporal resolution of PET and fMRI, it has been impossible to determine whether neuronal activity in motor and premotor regions was related to stimulus processing or to internal speech production (i.e., covert rehearsal) following stimulus processing. In the present study, this possibility was addressed via the temporal resolution of the EEG. If covert rehearsal or working memory for the response was responsible for activity in motor regions, such activity would not be expected prior to an acoustic stimulus. Further, based on models of speech production, a minimum of 200ms would be required for participants to process an auditory stimulus and to initiate covert rehearsal (Callan et al., 2010). As such, if

differences between correct and chance trials are found immediately following stimulus offset, it is unlikely that such differences would be due to covert rehearsal.

## Methods

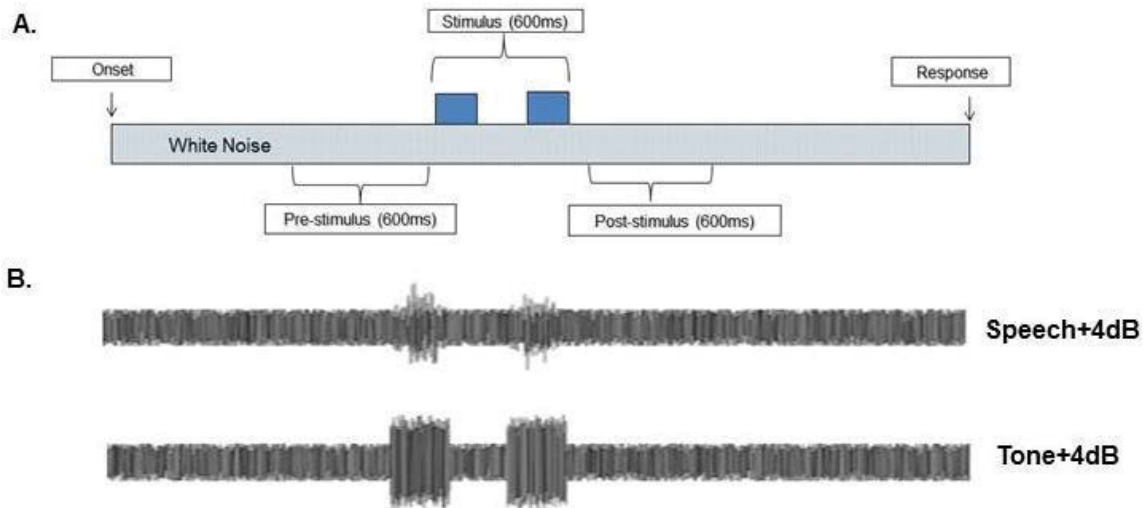
### Participants

Sixteen right-handed English-speaking adults (15 female and 1 male) with a mean age of 25 (range 20-42) participated in this study. Participants reported no diagnosed history of communicative, cognitive or attentional disorders. Degree of handedness was assessed using the Edinburg Handedness inventory (Oldfield, 1971). Prior to the experiment, informed consent approved by The University of Tennessee Health Science Center Institutional Review Board was obtained for all participants.

### Stimuli

Speech stimuli consisted of /ba/ and /da/ syllable generated using AT&T naturally speaking text-to-speech software. The software generates syllables using speech synthesized from a human male speaker. Half of the stimuli had different initial sounds (e.g., /ba/ and /da/) and the other half were the same (e.g., /ba/ and /ba/). The stimuli were normalized to have the same root-mean-square (RMS) amplitude and low-pass filtered with a cutoff at 5kHz. Each stimulus syllable was 200ms in duration with an interstimulus interval of equal length (i.e., 200ms). Thus, the total time required to present a stimulus pair was 600ms. For the tone discrimination task, sine-wave tone sweeps were generated using a procedure adapted from a previous neuroimaging study (Johannis & Gati, 2003). Tone-sweep stimuli were composed with an 80ms modulated tone onset and a 120ms steady state 1000Hz sine-wave. As for the speech stimuli, tone-sweeps were generated, low-pass filtered with a cut-off at 5kHz, and normalized to have the same RMS amplitude as the speech stimuli. Tone pairs differed only in whether the pitch onset was lower at 750Hz than the steady state tone or higher at 1250Hz. For both speech and tones the time between trials (i.e., interstimulus interval) was 3000ms. White noise for the tone and speech stimuli was generated and processed using the same procedure as for the speech sounds, with a low-pass filter cut-off at 5kHz. All auditory stimuli were processed using Soundtrack Pro academic software on an iMac (2 GHz intel core duo) computer and were sampled at 44kHz. Conditions were placed in random order prior to presentation. All stimuli were presented at an absolute intensity of 70dB. An example time line of one stimulus trial is displayed in Figure 3-1A along with examples of normalized speech and tone-sweep stimuli at a +4dB signal-to-noise ratio (SNR) as function of amplitude over time in Figure 3-1B.

Previous investigations have shown better than chance performance on a forced choice syllable discrimination task using a +4dB SNR and chance performance using a -6dB SNR (Binder et al., 2004; Callan et al., 2010). However, pure tones may be detected



**Figure 3-1. Sample time-line of one trial and normalized stimulus amplitude over time.**

Notes. A) time-periods of interest indicated prior to stimulus onset, during stimulus presentation, and following stimulus presentation and B) amplitude over time pre-stimulus, during stimulus, and following stimulus for both speech (top) and tone-sweep stimuli (bottom) for Actsp+4dB and Actn+4dB.

with noise intensities as high as 18dB above pure tone intensity (i.e., -18dB SNR) (Ernst, Verhey, & Uppenkamp, 2008). To account for differences between perceived loudness between tone and speech stimuli, preliminary behavioral data was collected on 10 female participants using Stim2 presentation software presented through Etyomotic ER1-14A tube phone inserts in a sound treated booth. Syllable and tone stimuli were embedded in white noise and presented in 20 trials using the time scheme as in Figure 3-1A at the following SNRs -18dB, -12dB, -6dB, +4dB. Syllable stimuli were identified above chance in the +4db condition only. Accuracy for tone- sweep conditions were not above chance in -18dB SNR, with 60% in -12dB SNR, 78% in the -6dB condition, and 75% in +4dB condition. Paired-t tests revealed no significant difference ( $p>.05$ ) between the +4dB and -6dB conditions. As such, the SNRs for speech were set at -6dB and +4dB and for tones at +4dB and -18dB.

## **Procedure**

Stimuli were presented using Stim 2 4.3.3 stimulus presentation software on a PC computer. The experiment was conducted in an electronically and magnetically shielded, double-walled, sound-treated booth. Participants were seated in a comfortable reclining armchair with their heads and necks well supported. Participants were told that they would be listening to white noise, syllables, and tones. They were instructed that the onset of one trial would commence when white noise was audible, followed by either syllable or tone stimuli. Participants were asked to indicate whether the syllables or tones sounded the same or different by pressing a button using the left thumb only. As premotor planning for finger and hand movements is generally brief (~200ms) (Callan et al., 2010; Grafton & Pfurtscheller, 2006), to further control for the possibility that preparation for the response might confound premotor activity related to stimulus processing, participants were signaled to respond via a 100ms, 1000Hz sine wave tone 1400ms after stimulus onset (see Figure 3-1A). To control for stimulus-response bias in the button press task, the order of the button press was counterbalanced.

All conditions were randomized prior to presentation and presented in two randomized blocks consisting of 40 trials each. Performance was evaluated as a percentage of correct trials (%CT). Participants were asked to listen under the following conditions: 1) Passively listening to noise (Pasn); 2) Passively listening to speech syllables in +4dB noise (Passp+4dB); 3) Passively listening to tone-sweeps in +4dB noise (Pastn+4dB); 4) Active syllable discrimination-in +4dB noise (Actsp+4dB) 5); Active tone-sweep discrimination-in +4dB noise (Actn+4dB); 6) Active syllable discrimination in -6dB noise (Actsp-6dB); 7) Active tone-sweep discrimination in -18dB noise (Actn-18dB).

## **Data Acquisition**

Thirty-two electrode channels were used to acquire EEG data based on the extended international 10-20 method of electrode placement (Jasper, 1958) using an

unlinked, sintered NeuroScan Quik Cap. Recording electrodes included Cz, C3, C4, CP4, CP3, Pz, P3, P4, P8, Fz, F3, F4, F7, F8, FC4, O1, O2, FP1, FP2, FT7, FT8, T3, T4, T5, T6, TP8, TP7 with two electrodes on the left (M1) and right mastoids (M2). The reference electrode was placed on the nasion and the ground electrode was at Fpz. The electro-oculogram (EOG) was recorded by electrodes placed on the left superior orbit and the left inferior orbit (VEOG) and on the lateral and medial canthi of the left eye (HEOG) to monitor vertical and horizontal eye movements, respectively. The impedances of all electrodes were measured at 30 Hz before, during, and after testing and were never greater than 5 K $\Omega$ .

EEG data were collected using Compumedics NeuroScan Scan 4.3.3 software and the Synamps 2 system. A mean of 50 trials (~480s) of artifact free data were obtained for each condition. The raw EEG data was filtered (0.15 – 100 Hz), and digitized via a 24-bit analog-to-digital converter at a sampling rate of 500 Hz. Data was time-locked to the onset of individual speech perception trials. After data collection, the recorded EEG signal and electro-oculogram (EOG) data was segmented into single trials lasting approximately 5000ms each, spanning from -3000ms to +2000ms with reference to stimulus onset (i.e., zero time). That is, time prior to syllable and tone-sweep stimuli was considered negative and time following syllable and tone-sweep stimuli was considered positive. To examine pre and post-stimulus activity, the EEG data was segmented (epoched) into the 3000 milliseconds pre-stimulus period, and 2000ms post stimulus onset. EEG data were visually inspected and trials contaminated by gross artifacts greater than 200  $\mu$ V were removed. A minimum contribution of 40 epochs for each participant in each condition was required for inclusion in the experiment. Due to a contribution of only 20 trials in several conditions, one participant was omitted from analysis.

## Data Analysis

### ICA Preprocessing

Prior to ICA training, EEG data were concatenated (i.e., added without mixing) for each participant across conditions. Subsequent ICA training was implemented using the extended *runica* algorithm implemented in EEGLAB. The initial learning rate was set to .001 with a stopping weight of 10<sup>-7</sup>. Linear decomposition using the extended Infomax algorithm (Lee, Girolami, & Sejnowski, 1999) was conducted for each participant across experimental conditions. The algorithm spheres or decorrelates the data matrix prior to ICA rotation and computes the variance of IC projection weights on to the original EEG channel data (Onton et al., 2005; Pfurtscheller et al., 2006). The resulting square weight matrix (30x30) and its inverse ( $W^{-1}$ ) back projected onto the original EEG channels are thus applied to each subject, yielding a single set of weights for each experimental condition expressing independence in the data. This process allows for a comparison of condition differences for the same set of component weights. Thus, for each subject all conditions bear the same set of component weights, allowing for a fair comparison



between conditions. The inverse weight matrix can then be back projected onto the original EEG channel configuration, providing spatial scalp topography for the components.

Independent components (IC's) were evaluated for each participant across experimental conditions using three criteria. First, an automated algorithm (ADJUST) shown in a previous study to have good inter-rater reliability with researchers experienced in IC noise removal, was used to tag non-brain artifact components in the EEGLAB software (Mognon, Jovicich, Bruzzone, & Buiatti, 2010). Scalp-maps and log spectra were also visually inspected for indicators of non-brain artifact including high power in low frequencies, abnormal spectral slope, and scalp-topographic distributions known to be associated with eye-movement and temporal muscle contraction (see Figures A-10 through A-12). ICs with characteristic signs of non-brain artifact were then pre-tagged for each subject. Second, IC's with 20 trials having outlier values ( $\mu\text{V}$  SD set to 10) over the electrode with maximum power were eliminated (Callan et al., 2010). Finally, equivalent current dipole (ECD) models for each component were computed using a four-shell spherical head model (BESA) in the DIPFIT toolbox (Oostenveld & Oostendorp, 2002, freely available at [scn.ucsd.edu/eeglab/dipfit.html](http://scn.ucsd.edu/eeglab/dipfit.html)). Standard electrode coordinates were warped to the BESA head model followed by coarse and fine-fitting to the spherical wire matrix, yielding dipole models for each of 480 ICs.

### **sLORETA Source Estimations**

sLORETA is a functional imaging technique that provides standardized linear solutions for modeling 3-D distributions of the likely cortical generators of EEG activity (Pascual-Marqui, 2002). The software uses a 3-D spherical head model separated into compartments including, the scalp, skull, and brain. sLORETA analysis operates under the assumption that scalp-recorded signals originate primarily in the cortical gray matter/hippocampi and that neighboring neurons are synchronously activated, giving rise to a signal that is distinct from surrounding noise. The head model is standardized with respect to the Talairach cortical probability brain atlas, digitized at the Montreal Neurological Institute (MNI) and uses EEG electrode coordinates derived from cross-registrations between spherical and realistic head geometry (Towle et al., 1993). The brain compartment includes 6239 voxels (5mm resolution). Electrode coordinates were converted to ASCII text format and exported to sLORETA from the EEGLAB module. For each IC, inverse ICA weight projections onto the original EEG channels were exported to the sLORETA data processing module for each participant. Cross-spectra were computed and mapped to the standard Talairach brain atlas cross-registered with the Montreal Neurological Institute (MNI) coordinates, yielding sLORETA estimates of current source density for each of 480 ICs.

## Independent Component Clustering

To identify similar independent components across participants, 480 (30x16) components were then clustered using measure product methods in the K-means statistical toolbox implemented in EEGLAB. The toolbox uses principle component clustering methods to reduce data dimensions and yields similar component clusters across participants. Product clustering considers precomputed measures with the aim of dividing  $M$  points by  $N$  dimensions into a specified number of clusters so that the within cluster sum of the squares is minimized. Here, 28 possible component clusters were considered. The data dimensions were reduced to 10 with the standard deviation set to 3. As such, ICs more than 3 standard deviations from any cluster mean were excluded as an outlying cluster. As the sensorimotor  $\mu$  rhythm is known to have characteristic peaks at  $\sim 10$  and  $\sim 20$ Hz and source locations within the sensorimotor cortex, scalp-maps, log spectra, and equivalent current dipole modes were precomputed and considered in the clustering analysis. Component power spectra for each subject were calculated by averaging fast Fourier transform (FFT) spectra for each epoch using a window length of 256 points. Scalp topographies were computed as 30 channel (x,y) map gradients and ECD models were precomputed in the manner described in a previous section.

After clustering, only components with a single dipole model within the head volume accounting for 80% or greater of the variance in the independent component scalp distribution were included in  $\mu$  component clusters. Pre-identified noise components tagged prior to the analysis were used to identify clusters accounting for non-brain sources. In two cases, cluster membership was adjusted considering both dipole location and sLORETA current source maximum and distribution. Particularly for ICs in which a single dipole model accounted for less of the variance percentage (i.e.,  $>15\%$  residual variance), the sLORETA distribution was used to investigate whether the distribution was within the sensorimotor cortex. Only dipole locations and sLORETA source maximum voxels within the precentral and postcentral gyri with spectral peaks near 10 and 20Hz were included in  $\mu$  component clusters.

## Event-related Spectral Perturbations

To examine stimulus induced changes in the EEG spectrum ERSP transforms were precomputed in the EEGLAB module using the STUDY command structure. ERSPs are changes scaled in normalized decibel units from a chosen spectral baseline over a broad spectral range (here 0-40Hz) (Delorme & Makeig, 2004). For independent components, ERSPs are scaled in RMS decibel units on the same scale as the component. Thus, IC scalp map topographies and ERSPs share the same RMS scale in decibel units. In this study, ERSPs were computed using a Morlet sinusoidal wavelet set at 3 cycles at 3Hz rising linearly to 20 cycles at 40Hz. Here, a 1000ms pre-stimulus baseline was selected from the silent interstimulus interval. This baseline served as a time period during which a surrogate distribution was generated. The surrogate data distribution is constructed by selecting spectral estimates for each trial from randomly selected latency windows in the specified epoch baseline. In this study, the baseline data was sampled 200

times, producing a baseline distribution whose percentiles may then be taken as significance thresholds (Delorme & Makeig, 2004). Significant changes in spectral power (i.e., increases or decreases from baseline) were then tested using a bootstrap resampling method. Significant differences from baseline ( $p < .05$  uncorrected) were considered in the subsequent within subjects analysis of ERSPs for identified  $\mu$  component clusters.

Analysis of condition effects for the left and right  $\mu$  ERSPs were carried out using the STUDY command structure in EEGLAB. The single trial current for all seven experimental conditions for frequencies from 0-40Hz and times from approximately -600ms to 2000ms post-stimulus onset were entered into a time-frequency analysis. For the two conditions in which performance was better than chance (Actsp+4dB and Actn+4dB) only trials discriminated correctly were considered in the ERSP analysis. A mean of 64 trials across conditions were entered into the ERSP analysis. Wavelet estimates across trials for each time and frequency were then converted to a time-frequency matrix (69x105) with the 3.4Hz to 39.9Hz to -589 to 1441ms. To test the significance of condition effects, non-parametric random permutation statistics in a 1x7 repeated measures ANOVA design were computed. The advantage of using non-parametric statistics for hypothesis testing of ERSPs is that this approach does not assume that the data are normally distributed. As discussed in previous papers, the event-related spectral increases (ERS) and decreases (ERD) that characterize ERSPs are frequently non-normal (Graiman & Pfurtscheller, 2006).

Random permutation methods generate a surrogate distribution constructed from time-frequency values randomly shuffled from each condition across all possible permutations. This random distribution represents the null hypothesis that no condition differences exist. In this study, 2000 random permutations were computed and compared to  $F$ -values for the mean condition differences. To control for the inflation of Type I error rates associated with multiple comparisons, a correction for false-discovery rate ( $p$ FDR) was applied, allowing for a conservative test of condition effects (Benjamini & Hochberg, 2000). The FDR adjusts the significance threshold using a false discovery rate computed from the random permutation tests.

## CHAPTER 4. RESULTS

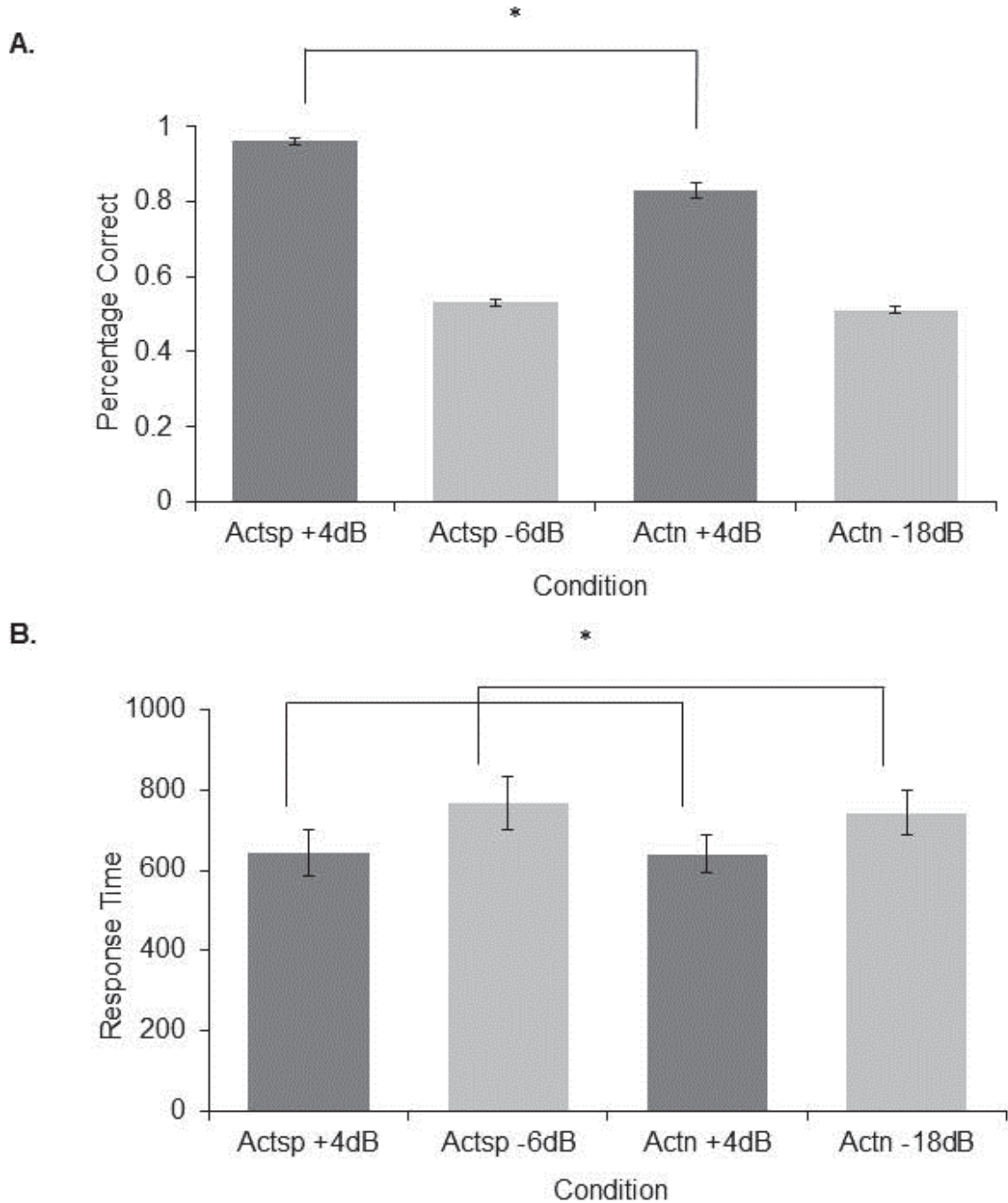
### Behavioral Performance

#### Percentage Correct

The means and standard errors for % correct trials (CT) in the four active speech and tone perception conditions are displayed in Figure 4-1A. Prior to the analysis trials with response times (RT) greater than three standard deviations from the mean response time (i.e., trials greater than 1996 milliseconds) were removed and were not considered in any subsequent analysis. Performance on the perceptual identification tasks (i.e., /ba/ vs /da/ and /1250Hz/ vs /750Hz/) was assessed as a percentage of correct trials. Performance in the Actsp+4dB condition was at near ceiling levels with a mean of 96% (SE= .01) correct. The Actsp+4dB condition was associated with better performance than the Actn+4dB condition, with a mean of 83% (SE=.02) correct. The mean for the Actsp-6dB and Actn-18dB were 53% (SE=.01) and 51% (.01) correct respectively. For the active conditions, a repeated measures analysis of variance (ANOVA) for the factor condition (1x4) revealed a significant main effect [ $F=(3)131.65, p<.01$ ]. A series of planned orthogonal single degrees of freedom comparisons were employed to determine condition differences. A significant difference was found for a comparison between % CT in the Actsp+4dB condition and the Actn+4dB condition [ $F=39, p<.01, \eta^2=.72, \Phi =1$ ]. No significant difference was found for a comparison of the Actsp-6dB and Actn-18dB conditions [ $F=1.79, p=.20$ ]. The Actsp-6dB and Actn-18dB were also not significantly different from chance level performance [ $t=.98, p=.20$ ].

#### Response Time

The means and standard errors for button press response time are depicted in Figure 4-1B. Response times (RT) were evaluated as the time in milliseconds from the cue to respond (i.e., 1000Hz tone) to the button press response. RTs for each subject in the four active conditions were entered into a repeated measures ANOVA with the factor condition (1x4). The analysis revealed a significant main effect for condition [ $F=3.71, p<.01, \eta^2=.19, \Phi =.77$ ]. A series of planned, single degrees of freedom a priori contrasts revealed significant differences between correct trials in the Actsp+4dB and Actn+4dB compared to chance level performance trials in the Actsp-6dB and Actn-18dB conditions respectively [ $F=7.23, p<.01, \eta^2=.32, \Phi =.71$ ]. No significant difference was found between the Actsp+4dB and Actn+4dB conditions [ $F=.00, p=.96$ ] or between the Actsp-6dB and Actn-18dB [ $F=.24, p=.62$ ]. The mean RT for the two conditions in which performance (Actsp+4dB and Actn+4dB) was above chance were 642ms (SE=58) and 641ms (SE=47) respectively. The mean RT for the two conditions in which performance was at chance levels was 767 (SE=68) and 743ms (SE=55). Taken together, the analysis revealed an inverse relationship between perceptual performance in the active conditions and button press response time.



**Figure 4-1. Means and standard errors for percentage correct trials and response time.**

Notes. A) active speech and tone perception conditions and B) active speech and tone perception conditions. Significant condition differences at  $p < .05$  are indicated by \*.

## Electrophysiological Measures

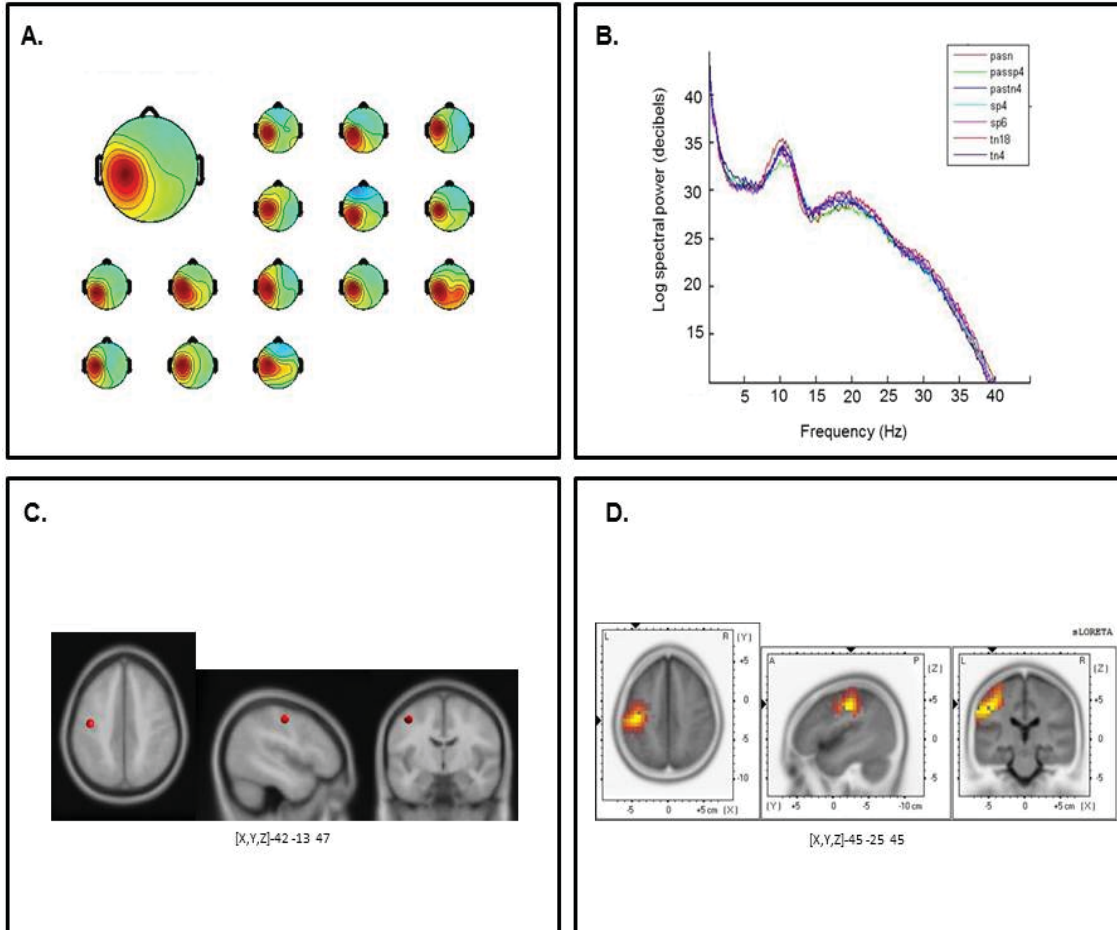
### Independent Component Clustering

Independent component clustering revealed six distinct component clusters with likely neural as opposed to non-brain (i.e., artifact) sources. Average log spectra, ERSPs, scalp topographies, and source estimates are presented in the Appendix (Figures A-1 through A-13). Six component clusters accounted for eye-blinks and vertical eye-movements, horizontal eye-movements, temporal muscle noise, and nonspecific (electromagnetic line-noise). Sample wave forms, log spectra, and scalp topographies of blinks, eye-movements, and temporal muscle noise are presented in the Appendix (Figures A-11 through A-13). Component clusters with similar scalp-topographies, spectra, ECD, and sLORETA CSD locations were found for a bilateral frontal midline cluster (Figure A-7), central midline cluster (Figure A-8), left and right posterior temporal clusters (Figures A-5 and A-6). The frontal and central midline components showed relatively high spectral power in the delta and theta range (0-5Hz). The temporal components were consistent with previous evidence finding temporal components with peak spectra at ~10Hz in a speech perception task (Callan, Callan, Kroos, & Vatikiotis-Bateson, 2001). However, because the focus of the current investigation is on the sensorimotor  $\mu$  rhythm, only left and right sensorimotor components are discussed further.

Fourteen participants submitted ICs with the hallmark characteristics of the left sensorimotor  $\mu$  rhythm and 13 participants submitted ICs with hallmarks of the right  $\mu$  rhythm. The left cluster revealed mean scalp-topographies centered over the left sensorimotor cortex (Figure 4-2A) whereas the right cluster showed a similar topography over the right hemisphere (Figure 4-3A). For both clusters, log spectra revealed distinct spectral peaks at ~10Hz and ~20Hz (Figures 4-2B and 4-3B) and ECD locations within the left and right pre or postcentral gyri (Tables 4-1 and 4-2 ) with an average dipole location at Talairach coordinates [(x,y,z) -42,-13, 47] in the left hemisphere and [(x,y,z) 41,-12,42] in the right hemisphere.

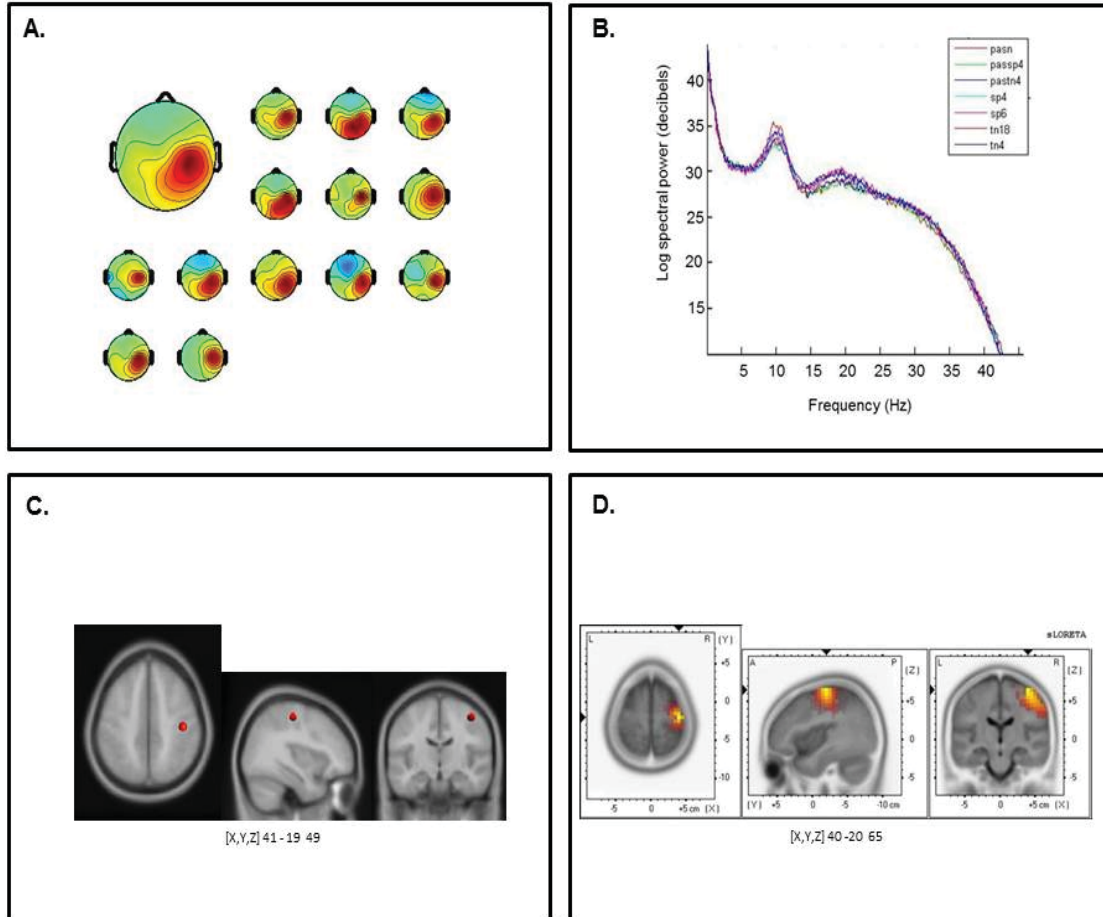
To evaluate the statistical significance of dipole locations across participants, statistical comparisons relative to zero (i.e., no activation) were computed for all  $\mu$  scalp topographies in the sLORETA statistical module (Grin-Yatsenko, Baas, Ponomarev, & Kropotov, 2010). A single paired t-test was carried out for each frequency between 0-40Hz (159 frames) with the smoothing parameter set to 1 (single common variance for all variables), and 5000 random permutations yielding corrected t-value thresholds for all 6,235 voxels in the sLORETA solution space. The paired test revealed significant voxels at  $p < .001$  (i.e., dipole locations) in the precentral and postcentral gyri with maximum current source density estimates at Talairach [(x,y,z) -45,-25, 45] in the left hemisphere and Talairach [(x,y,z) 40,-20,55] in the right. To determine the extent to which sLORETA and ECD estimates were related, a bivariate correlation analysis was conducted on the maximum CSD coordinates and the ECD coordinates for each sensorimotor IC. The analysis revealed correlation coefficient of  $r = .93$  that was





**Figure 4-2. Cluster results for the left-hemisphere  $\mu$  component.**

Notes. A) mean scalp potential distribution ( $W^{-1}$ ) scaled to RMS microvolts, B) mean spectra of the component as a function of condition in relative decibels, C) average equivalent current dipole location, and D) maximum current source density voxels ( $p < .001$  corrected for multiple comparisons).



**Figure 4-3. Cluster results for the right-hemisphere  $\mu$  component.**

Notes. A) mean scalp potential distribution ( $W^{-1}$ ) scaled to RMS microvolts, B) mean spectra of the component as a function of condition in relative decibels, C) average equivalent current dipole location, and D) maximum current source density voxels ( $p < .001$  corrected).



**Table 4-1. Percentage residual variance for each ECD model, ECD x,y,z coordinates, current source density x,y,z coordinates, and component rank for the left hemisphere  $\mu$  cluster.**

<b>Participant</b>	<b>% RV</b>	<b>ECD <math>L\mu</math> (x,y,z)</b>	<b>CSD <math>L\mu</math> (x,y,z)</b>	<b>Rank</b>
1	11.16	-51 -22 42	-55 -30 45	22
2	1.78	-31 -3 32	-35 0 20	15
3	.79	-33 1 49	-35 -3 42	19
4	2.21	-25 -26 61	-25 -26 61	16
5	3.89	-52 -19 31	-54 -23 33	17
6	2.06	-37 -4 78	-40 -12 60	22
7	2.62	-43 -22 37	-45 -22 43	22
8	2.74	-32 -7 44	-30 -3 46	12
9	1.74	-32 -19 40	-35 -22 38	16
10	15.44	-31 18 62	-30 12 45	14
11	.71	-41 -16 52	-45 -17 47	18
12	11.92	-37 -8 30	-54 -18 29	15
13	3.15	-42 -11 47	-40 -12 47	13
14	1.41	-56 -20 19	-54 -18 19	13

**Table 4-2. Percentage residual variance for each ECD model, ECD x,y,z coordinates, current source density x,y,z coordinates, and component rank for the right hemisphere  $\mu$  cluster.**

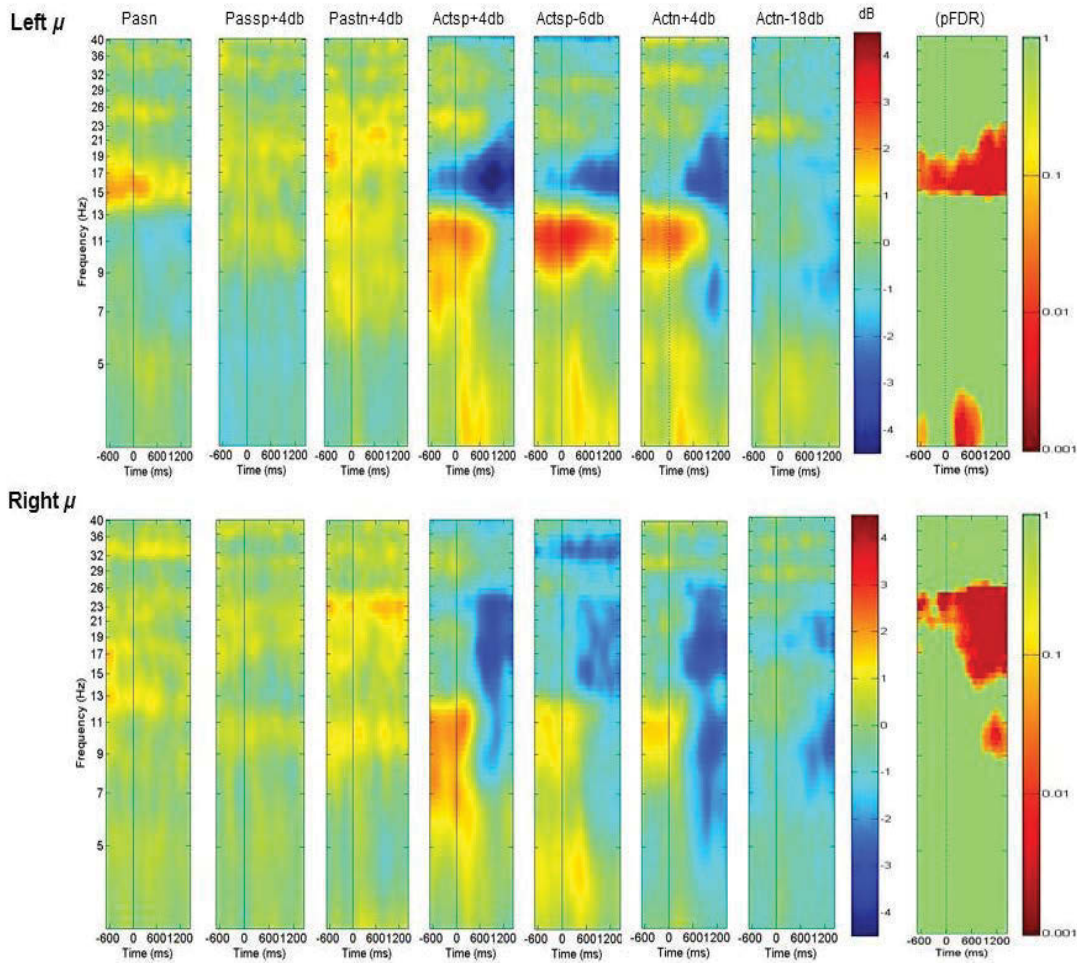
<b>Participant</b>	<b>% RV</b>	<b>ECD <math>R\mu</math> (x,y,z)</b>	<b>CSD <math>R\mu</math> (x,y,z)</b>	<b>Rank</b>
1	10.09	45 -20 53	46 -28 52	17
2	6.50	40 -18 46	45 -22 47	16
3	5.74	37 -15 46	37 -11 65	6
4	1.78	31 -1 43	30 -3 46	12
5	14.22	40 -25 54	59 -27 47	21
6	2.58	31 -8 58	30 -11 65	13
7	2.73	44 -43 30	59 -55 29	13
8	1.78	35 -11 60	35 -11 65	16
9	7.64	57 -22 63	50 -26 57	23
10	3.64	45 -20 58	45 -16 61	11
11	.80	55 -14 35	50 -18 38	16
12	11.99	58 -21 56	54 -17 52	20
13	2.73	44 -43 30	59 -52 35	13

significant at  $p < .01$ , suggesting a close relationship between CSD and ECD estimates of IC source locations.

### **Event-related Spectral Perturbations: Left $\mu$**

A repeated measures ANOVA with the factor condition (1x7) revealed no significant differences for the number of trials submitted between conditions ( $F = .92$ ,  $p = .48$ ). In the STUDY command structure, mean ERSP values across subjects and conditions are shown in a time-frequency scalogram with corrected significance values for condition depicted in a separate figure. Non-significant values are depicted in green and significant values are depicted in color from orange for weaker values to red for stronger values ( $pFDR < .10$  to  $pFDR < .001$ ). The initial permutation analysis (1x7) revealed significant ERSPs at  $pFDR < .05$  in the 15-20Hz range (beta) and the 0-4Hz range (delta) for the left  $\mu$  component (see Figure 4-4). Significant time-frequency values were found in the time-periods prior to, during, and after stimulus onset with a peak event-related decreases in spectral power (i.e., ERD) in the time period after stimulus offset. To determine the sources of condition effects, two separate ANOVA designs were computed using the STUDY command structure. Because the time periods before, during, and after stimulus onset were of interest, all subsequent analysis were restricted to the equal 600ms time intervals prior to, during, and following stimulus onset (i.e., -600-1200ms).

First, to determine whether any significant differences existed between the Pasn baseline and the other passive conditions, a 1x3 ANOVA was conducted for the Pasn, Passp+4dB, and Pastn+4dB conditions. No significant differences corrected across the entire time-frequency matrix (69x92) were found ( $pFDR > .05$ ) in either the 15-20Hz range or the 0-4Hz range, indicating no differences between the Pasn baseline condition and the other two passive conditions. To assess which conditions were significantly different from the Pasn baseline, a series of paired t-tests were performed. Significant differences ( $pFDR < .05$ ; 69x92) for the time periods before, during, and after stimulus onset were found for the Actsp+4dB and Actsp-6dB only (Figure A-1). Analysis of the active conditions in which discrimination was required (1x4; Actsp+4dB, Actsp-6dB, Actn+4dB, and Actn-18dB), revealed a significant main effect ( $pFDR < .05$ ; 8x92) in the 15-19Hz range for the time period between 600-1200ms following stimulus offset (Figure A-3). A comparison between the Act+4dB and Actsp-6dB conditions across the 15-19Hz range between 600-1200ms period following stimulus offset ( $pFDR < .05$ ; 6x15) revealed a significantly larger peak ERD in the Actsp+4dB condition. As such, the left component cluster showed significant effects for only the syllable discrimination task and further showed significant differences in the time period following stimulus offset for correct discrimination trials in the Actsp+4dB condition relative to the chance trials in the Actsp-6dB condition.



**Figure 4-4. Mean left and right hemisphere  $\mu$  time-frequency ERSPs (event-related spectral perturbations).**

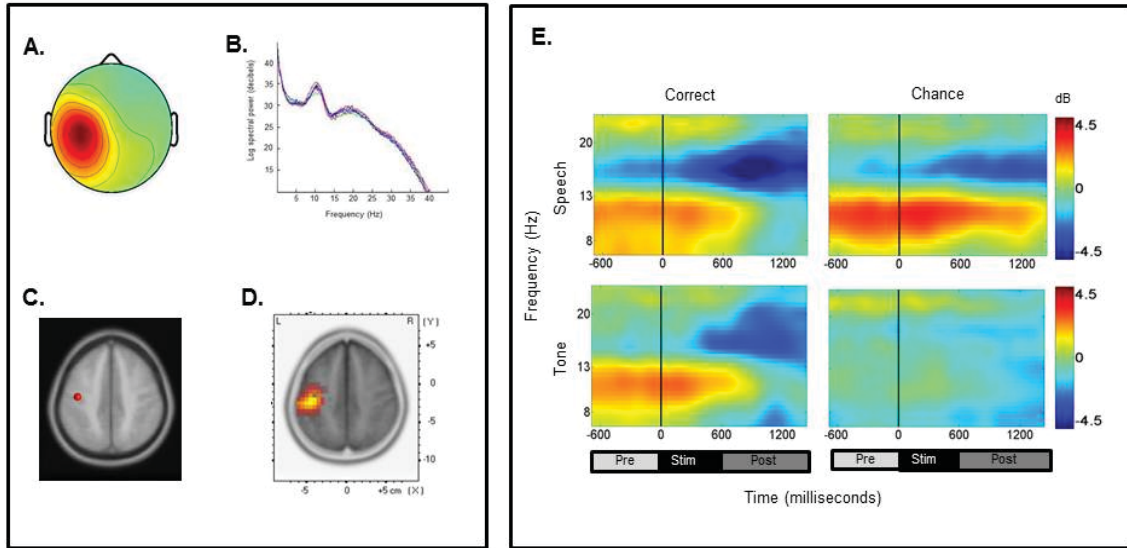
Notes. ERSPs are scaled in the same root-mean-square decibel units as a function of condition (1x7) and random effects analysis indicating significant values in the traditional beta, alpha, and delta ranges. Non-significant values are colored green, with significant values shown in orange and red. Event-related decreases in spectral power are indicated in blue (-4.5) and increases are indicated in red (4.5).

## Event-related Spectral Perturbations: Right $\mu$

The initial permutation analysis revealed significant ERSPs at  $pFDR < .05$  in the 15-25Hz range (beta) for the right  $\mu$  (Figure 4-4) component. Significant time-frequency values corrected across the entire time-frequency matrix ( $pFDR < .05$ ;  $69 \times 105$ ) were found in the time-periods prior to, during, and after stimulus onset with a peak event-related decrease in spectral power in the time period after stimulus onset. To determine the sources of condition effects, a 1x3 design for the passive conditions (Pasn, Passp+4dB, and Pastn+4dB conditions) was conducted. The ANOVA revealed no significant differences ( $69 \times 92$ ;  $pFDR > .05$ ). To assess which conditions were significantly different from the Pasn condition (i.e., the baseline), a series of paired contrasts were performed. Significant differences ( $pFDR < .05$ ) for the time periods before, during, and after stimulus onset were found for correct trials in the Act+4dB and chance trials in the Actsp-6dB conditions (Figure A-2). Significant spectral suppression from baseline ( $pFDR < .05$ ) in the active tone discrimination conditions (Actn-4dB and Actn-18dB) were found for the time period following stimulus onset only. Although ERDs were found in the time-period prior to tone discrimination trials, they did not fall below the significance threshold of  $pFDR < .05$  (see Figure A-3). Analysis of the active conditions in which a sensory-decision was required (i.e., Actsp+4dB, Actsp-6dB, Actn+4dB, and Actn-18dB), revealed no significant differences in 15-25Hz range ( $pFDR < .05$ ;  $69 \times 92$ ). Thus, although active tone discrimination conditions differed from the passive noise baseline, no significant differences were noted between the active conditions.

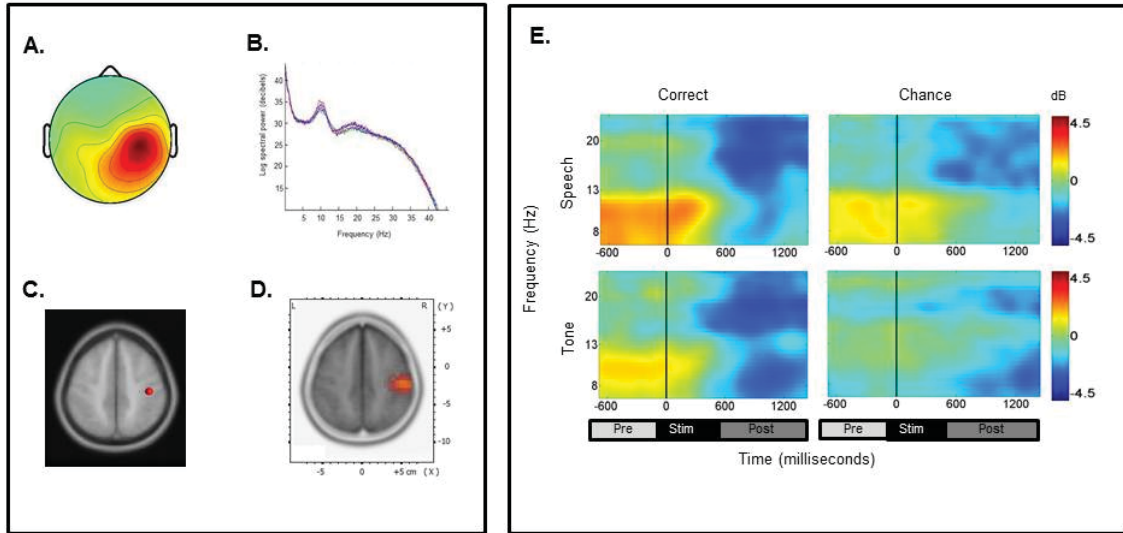
Previous studies have suggested that reaction time may provide a measure of sensory decision processes related to a response. To investigate the possibility that ERSPs following stimulus onset might be related to the button press response, significant ERSPs values (i.e. 15-20Hz for the left  $\mu$  and 15-25Hz for the right  $\mu$ ) were extracted from correct trials in the +4dB condition and averaged over three time-periods of interest prior to, during, and following stimulus onset (-600-0, 0-600, and 600-1200). A bivariate correlation analysis was performed on each participant's mean ERSP values in the time period following stimulus offset (600-1200) and RT for each participant. The results indicated no significant correlation between RT and ERSPs ( $r = .02$   $p = .94$ ).

In summary, left and right  $\mu$  rhythm clusters were associated with suppression relative to Pasn during the active syllable discrimination task prior to, during, and following the onset of syllable stimuli. Relative to chance trials, correct level trials were associated with significantly greater suppression in the time period immediately following acoustic syllable stimuli only (see Figures 4-5 and 4-6). By contrast, tone discrimination trials were not associated with suppression prior to stimulus onset relative to Pasn. Tone trials were significantly different from Pasn in the time period after stimulus presentation only. Finally, no significant differences were found between correct tone discrimination trials and trials discriminated at chance levels.



**Figure 4-5. Measures for the left-hemisphere  $\mu$  cluster.**

Notes. A) mean scalp-topography ( $W^{-1}$ ), B) mean dipole location, C) mean spectral power across conditions, D) sLORETA solutions, and E) mean time-frequency ERSPs (event-related spectral perturbations) as a function of stimulus type (speech and tone) and performance level (correct and chance) for the time-periods prior to stimulus onset, during stimulus presentation, and after stimulus-offset.



**Figure 4-6. Measures for the right-hemisphere  $\mu$  cluster.**

Notes. A) mean scalp-topography ( $W^{-1}$ ), B) mean dipole location, C) mean spectral power across conditions, D) sLORETA solutions ( $p < .001$  corrected), and E) mean time-frequency ERSPs (event-related spectral perturbations) as a function of stimulus type (speech and tone) and performance level (correct and chance) for the time-periods prior to stimulus onset, during stimulus presentation, and after stimulus-onset.



## CHAPTER 5. DISCUSSION

To investigate the time-course of sensorimotor integration in speech and non-speech processing, the current study employed independent component analysis of event-related EEG to measure activity of sensorimotor  $\mu$  rhythm. To my knowledge, this is the first study to employ BSS methods to investigate suppression of the  $\mu$  rhythm in speech and non-speech discrimination task. As such, the present findings have important implications for the measurement of EEG during auditory tasks. Further, because performance levels on both tasks were manipulated, the present findings also have important implications for the relationship between sensorimotor integration and perceptual accuracy. Finally, as this is the only study to date to use non-invasive, low-cost EEG to measure the time-course of sensorimotor activity, findings may have implications for potential use as a brain computer interface (BCI) approach. Due to the complexity of the results, each of the specific aims, questions, and hypotheses will be addressed separately and subsequently framed within an overall discussion of the theoretical, methodological, and clinical significance.

### **Aim 1: To Identify Spatially Fixed and Temporally Independent Components Associated with Speech and Non-speech Processing**

The first aim of the current study was to apply independent component analysis to raw, linearly mixed EEG scalp-channel data to reveal maximally independent and spatially fixed EEG components associated with speech and non-speech processing. In the current study, cued speech and tone discrimination tasks in which participants anticipated upcoming percepts were employed. The application of the *runica* algorithm to the single-trial, concatenated epochs revealed three classes of independent components with bilateral distributions and shared features. The first class consisted of ICs with frontal and midline topographic distributions characterized by high theta band power and associated dipole models in the cingulate cortex spreading into the SMA. A second class of ICs with topographic distributions over the posterior temporal lobe were characterized by dipole models near primary auditory cortex and superior temporal gyrus with peak spectral power at  $\sim 10$ Hz. A third class of clusters was identified as the central sensorimotor or  $\mu$  rhythm. ICs in the sensorimotor cluster were associated with dipole models and sLORETA solutions in the sensorimotor cortex and showed two characteristic, distinct peaks at  $\sim 10$ Hz and  $\sim 20$ Hz. For all clusters, mean dipole models accounted for greater than 90% of the variance in the mean scalp-topographies, suggesting that those models provided a good fit for the associated ICs. The distributed sLORETA maximum CSD for the scalp-topographies was correlated with dipole-models for the same ICs, suggesting a high level of agreement between the two source solutions.

ICA findings in the present study lend further support to the notion that neural generators of the EEG give rise to signals with independent or nearly independent time courses associated with stimulus specific processing. Scalp-recorded EEG signals are thought to be generated by field potentials associated with a three centimeter squared

(~3Cm<sup>2</sup>) areas of cortex (Makeig et al., 2004; Onton & Makeig 2006). Because neuronal activity is weighted toward local connections (as opposed to the sparse interconnection between sources) and thus local synchronicity, the output of these small patches of cortex may be recorded as signals whose time-courses are independent or nearly independent of one another. The activity of the traditional EEG, including posterior alpha activity, sensorimotor rhythms, frontal, and central theta rhythms are thought to be generated by these locally weighted neuronal populations.

Previous studies using low and high-density electrode arrays have consistently identified ICs with the spectral signatures and topographic distributions consistent with known features of the traditional EEG and yield between 5 and 15 neurally generated components (Onton & Makeig, 2006). Decomposition of 32 channel electrode arrays has resulted in the identification of posterior and temporal alpha, frontal-midline theta band sources, and bilateral sensorimotor sources involved in a number cognitive tasks (Delorme et al., 2006; Makeig et al., 2004; Makeig et al., 2002a; Onton & Makeig, 2006) that may be modeled as physiologically simple sources (e.g., single dipole source) or small areas of cortex (e.g., distributed source solutions). Whereas these source generators are temporally independent, it is thought that cortical synchronization between generators may function to facilitate top-down and bottom-up communication both within and between brain regions. This large-scale interregional and interhemispheric communication is thought to be particularly important for anticipated perceptual signals as in the current paradigm.

As predicted in dual-stream theories, dipole models and sLORETA CSD solutions for IC clusters in the present study are consistent with a dorsal stream network known to be involved in active sublexical speech discrimination tasks (Hickok & Poeppel, 2004; 2007; Price et al., 2010; Turkeltaub & Coslett, 2010), including the SMA, M1, vPMC, posterior STG/Spt. Component clusters were also broadly consistent with two other studies employing ICA during auditory processing. Marco-Pallarés, Grau and Ruffini, (2005) found ICs in an auditory mismatch negativity (MMN) paradigm in frontal, cingulate, and superior temporal regions. A single subject EEG study of visual and auditory speech perception found similar independent component activations, with sLORETA distributions suggesting activation in the SMA, premotor and primary motor cortices, and in posterior temporal regions (Callan et al., 2001). Collectively, these regions are thought to be a part of the dorsal stream network involved in audio-motor transformation specifying 'how' phonemes were produced. The dorsal stream has also been implicated in a number of sublexical speech perception studies (Price et al., 2010; Turkeltaub et al., 2010) and is consistent with bottom-up and top-down anatomical connections between sensory and motor areas thought to make important contributions to the perception of speech (Davis & Johnsrude, 2007). As such, the most parsimonious explanation for the current results appears to be that cortical regions known to be active in auditory processing gave rise to maximally separated neuronal signals associated with small cortical areas within in the dorsal stream network.

## **Aim 2: To Identify Independent Components Associated with Known Features of the Sensorimotor $\mu$ Rhythm**

The second aim of the investigation was to identify ICs with hallmark features of the sensorimotor  $\mu$  rhythm. Several findings for left and right hemisphere components are consistent with features of the  $\mu$  rhythm. First, cluster analysis revealed left and right hemisphere IC clusters with topographic distributions over the sensorimotor cortex with spectral peaks at  $\sim 10\text{Hz}$  and  $\sim 20\text{Hz}$  and source locations in the lateral portion of the precentral gyrus. Second, a distributed localization approach indicated significant activations in the precentral and post central gyri, consistent with distributed source locations over the sensorimotor cortex within a small cortical area. Third, relative to the passive noise baseline, the left and right sensorimotor  $\mu$  rhythms in the active discrimination conditions were associated with differential suppression in the traditional beta range (here 15-25Hz), consistent with the activity of neuronal populations in the precentral gyrus (Hari, 2006). These findings strongly suggest that component clusters with the hallmark characteristics of the sensorimotor  $\mu$  rhythm were differentially involved in the processing of speech and non-speech auditory signals.

Although the  $\sim 10\text{Hz}$  and  $\sim 20\text{Hz}$  components of the  $\mu$  are thought to be phase locked and thus interdependent, there is reason to suspect functional differences between them. High-density MEG studies have shown power changes at 10Hz associated with dipole locations near the somatosensory cortex bilaterally clustering near the hand cortical region (Hari, Levänen, & Raij, 2000; Hari Salmelin, Mäkelä, Salenius, & Helle, 1997). The  $\sim 10\text{Hz}$  rhythm is also associated with tactile stimulation and binocular rivalry (Vanni, Rockstroh, & Hari, 1996), suggesting that it is more generally associated with somatosensory activity. Band pass filtering of the 10-12Hz upper alpha rhythm has been associated with suppression in a wide range of tasks, including auditory memory tasks (e.g., Sternberg test; Krause, 2006), and attention tasks in both the auditory and visual modalities (Klimesch et al., 1998), suggesting that the  $\sim 10\text{Hz}$  rhythm may not be specific to the processing of movement. Thus, the most likely role for the 10Hz rhythm during movement and movement processing is in coding the somatosensory consequences of the perceived or performed movement (Hari, 2006).

Several lines of evidence suggest that beta suppression (i.e.,  $\sim 20\text{Hz}$ ) over the sensorimotor region is generated primarily in the precentral gyrus. First, high-density MEG studies have shown that beta suppression during the overt production, imagination, and observation of movement is associated with dipole locations near the primary motor cortex bilaterally (BA4) following the moving body part (e.g., more lateral locations for mouth movements) (Hari, 2006; Hari et al., 1997; Simões, Salenius, & Curio, 2004). Second, in agreement with a source in the primary motor cortex,  $\sim 20\text{Hz}$  suppression has also been obtained from intracranial recordings within the central sulcus (Jasper & Penfield, 1949). Third,  $\sim 20\text{Hz}$  suppression is coherent with motor unit firing (Conway et al., 1995) and is enhanced by the administration of benzodiazapines, well known to result in clumsy, poorly controlled movements (Jensen et al., 2005). Finally, beta suppression has been shown to be inversely correlated with BOLD increases in M1 and the PMC during the observation and performance of movement (Yuan et al., 2010),

suggesting convergence between hemodynamic approaches with higher spatial resolution and electrophysiological approaches relying on local neuronal synchronization.

The functional distinction between the ~10Hz somatosensory and ~20Hz motor component in the current study is consistent with previous neuroimaging studies finding activity in the precentral gyrus (Skipper et al., 2005b; Wilson et al., 2004). Neuroimaging approaches have implicated overlapping regions within the precentral gyrus for both speech production and speech perception (Callan et al., 2010; Hickok et al., 2004; 2007; 2011; Wilson et al., 2004). In a high-density MEG study, Pullvermuller et al. (2007) found that ERPs occurred in 20ms steps during speech perception, moving from initial activity in the STG, to the IFG and M1 (i.e., primary motor cortex). Wilson et al. (2004) found overlapping peaks of BOLD activity in regions extending from the precentral gyrus (MNI  $x,y,z$ = -51, -11, 46) to the posterior bank of the central sulcus (MNI  $x,y,z$ =-45, 13, 34). As in the current study, for most subjects speech perception tasks were associated with increased activity in the precentral gyrus bilaterally. In accord with those findings, approaches employing both functional localization and TMS have found peak BOLD voxels in a similar region of the left superior PMC during speech production. With rTMS applied to the PMC and M1 region, modulations in speech discrimination, segmentation, and identification performance have been demonstrated (D'Ausilio et al., 2010; Meister et al., 2007; Sato et al., 2009).

More recently, Callan et al. (2010) found that a region within the precentral gyrus (vPMC) was more active (i.e., higher BOLD signal) for correct speech perception relative to incorrect speech perception, strongly implying that the motor system was involved in perceptual performance. In addition, using a speech morphing procedure in which speech percepts gradually emerged from a background of white-noise, Osnes, Hugdahl, Hjelmervik and Specht (2012) found that an area within the precentral gyrus (vPMC) was more active when speech percepts were identifiable as speech. Further, a recent activation likelihood meta-analysis of sublexical speech perception, showed a region of overlap between studies with maximal activation in a region near the vPMC, slightly more anterior than activations found in single studies (Turkeltaub & Coslett, 2010). In sum, a growing body of neurophysiological evidence has implicated the premotor and primary motor cortex within the precentral gyrus in the processing of speech in the auditory modality. Although the postcentral gyrus is known to be active during the observation of visual mouth movements, it is generally not found to be differentially active during speech processing in the auditory modality (Skipper et al., 2005b). As such, the current finding of differential beta suppression localized to the sensorimotor cortex is consistent with both known features of the  $\mu$  rhythm and activation of the precentral gyrus.

### **Aim 3: To Investigate Differential $\mu$ Rhythm Suppression Relative to Stimulus Onset, Stimulus Type, and Discriminability**

A third aim of the present study was to investigate whether suppression of the  $\mu$  rhythm differed as a function of stimulus onset, stimulus type, and task performance. This aim is significant not only for supporting theoretical models of speech processing but also

for addressing limiting factors associated with traditional neuroimaging methodology. A number of findings in this study are consistent with differential sensorimotor involvement specific to the accurate performance of a syllable discrimination task. First, in accordance with expectations behavioral measures indicated that active task performance differed as a function of SNR. A higher %CT and shorter RT in high SNR speech and tone conditions (Actsp+4dB and Actn+4dB) was found relative to low SNR speech and tone conditions (Actsp-6dB and Actn-18dB). As expected, due to the drastically decreased quality of auditory information in the low SNR conditions, perceptual performance was not above chance levels. Consistent with a previous investigations of speech-in-noise (Binder et al., 2004; Callan et al., 2010), increases in RT were associated with decreases in accuracy. This finding is consistent with the notion that continuous sensory decision processes require greater time when a given decision is ambiguous relative to when it is more easily identified (Binder et al., 2004). As such, it seems likely that sensory decision processes here may have required more time in the low SNR conditions, suggesting greater time to integrate time varying information involved in auditory processing with ongoing sensory decision processes.

However, most studies of speech-in-noise have assumed that decision mechanisms continue to operate until the point in time at which the results of sensory analysis meet the criteria required for a perceptual decision (Binder et al., 2004). In the present study, RT was operationally defined as the time required from the cue to respond until the point in time at which the participant made a decision. That is, participants were asked not to respond until they were cued via a tone to register a response. As such, it is likely that responses in this study required additional attentional and working memory resources not required in some previous studies. Considering internal models of speech processing, one possible alternative explanation for the well-known inverse relationship between accuracy and decision processes may be that participants test a greater number of alternative phonetic hypotheses about the acoustic signal as the ambiguity of acoustic cues increases (Skipper et al., 2006). This greater number of alternative hypotheses creates a delay in RT as early hypotheses are integrated with incoming sensory information.

Second, although performance levels were similar for both the speech and tone discrimination tasks, suppression of the  $\mu$  rhythm differed as a function of performance level for the syllable discrimination task only. Early activity (i.e., prior to stimulus onset) was present bilaterally for speech stimuli regardless of subsequent perceptual performance, suggesting precentral activation prior to stimulus onset consistent with internal model concepts. Third, brief peak activity (i.e., suppression) was significantly larger for the syllable discrimination task only when compared to extremely degraded stimuli (i.e., chance trials), suggesting that although the precentral gyrus is active for both types of stimuli it is greater when auditory features are sufficient to specify phonemic units. Fourth, significant suppression of the sensorimotor rhythm in the right hemisphere occurred for control tone stimuli in the time period following stimulus onset only. Thus, although speech stimuli elicit early activity peaking after stimulus discrimination, non-speech auditory stimuli are associated with activity following stimulus offset only and in the right hemisphere only. As such, early suppression for the



syllable discrimination task cannot be explained as a general attentional mechanism. Further, as no differences were found between correct and chance tone stimuli, it is unlikely that right precentral activation was important for task performance. Thus, it would appear that internal models were recruited for perceptual analysis in the syllable discrimination tasks only.

In addition to addressing theoretical predictions regarding the time-course of sensorimotor activity relative to acoustic stimuli, the experimental paradigm and findings address several methodological concerns that have traditionally limited findings of sensorimotor activity in classical neuroimaging studies. These limitations include contamination related to the response and covert rehearsal (Callan et al., 2010; Hickok & Poeppel, 2004; 2007). First, significant differences for stimulus type and performance level for tasks in which an active decision was required were found. As the button press and sensory decision requirements were the same for correct speech and tone trials, no significant difference in suppression or RT would be expected for the two types of stimuli if button press or sensory decision mechanisms accounted for differences. For example, in the left hemisphere, beta suppression was significantly different from the passive noise baseline for correct trials in the Actsp+4dB condition yet no significant suppression was observed for the Actn+4dB condition. However, as no difference in RT was observed between the two conditions, it is unlikely that sensory decision processes can account for suppression in the Actsp+4dB condition.

Second, previous studies have suggested that RT is a correlate of sensory decision processes. In the present study, no correlation between activity following stimulus onset and RT was observed suggesting that sensory decision mechanisms or working memory for a button press response cannot account for differential findings of  $\mu$  suppression. Third, previous studies have shown that manual movements are associated with greater power suppression in the hemisphere contralateral to the movement (Graumann & Pfurtscheller, 2006; Makeig et al., 2004). Here, the button press response was performed with the left thumb and suppression in the time period following stimulus offset was not significantly greater in the right hemisphere, strongly suggesting that the manual response cannot account for differential findings of beta suppression between correct and chance trials. Additionally, both the left and right rhythms peaked and then enhanced prior to the button press response, further indicating that the button press response cannot account for task related suppression (see Figure A-10).

Another way in which button press and sensory-decision processes were addressed was via the temporal resolution of the EEG. In the current experimental paradigm, the button press was delayed so that it occurred 1400ms after stimulus onset. Because peak suppression occurred 700-1100ms after stimulus onset (400ms prior to the response), it is unlikely to have been related to the button press or sensory-decision mechanisms as opposed to stimulus processing. In addition, significant beta suppression was found prior to and during stimulus processing for speech stimuli relative to baseline. Since suppression prior to a response or decision was found, it is highly unlikely that beta suppression in those time periods was related to either process.

In addition to memory and decision processes for response selection, it has been suggested that activity in motor regions may be due to internal speech production or covert rehearsal mechanisms functioning to enhance working memory (Hickok & Poeppel, 2007). Several findings in the current study are inconsistent with covert rehearsal. First, significant beta suppression of the  $\mu$  rhythm was found prior to stimulus onset for both speech conditions (i.e., correct and chance trials). As activity occurred prior to stimulus onset it cannot be related to covert rehearsal. Second, although it is possible that peak activity occurring after stimulus onset was related to internal production, because significant differences between correct and chance trials began less than 200ms (at  $\sim$ 100ms) following stimulus offset and lasted only 400ms, it is unlikely that this early activity was due to covert production. Third, since activity occurred continuously from -600ms prior to stimulus onset and peaked just after stimulus offset, it is unlikely that continuous beta suppression was due to two functionally distinct processes. Peak suppression for both speech and tone stimuli was brief, occurring within a 400ms span of time and diminishing thereafter (see Figure A-10), further suggesting that only one functional process was at work. If covert speech production for holding percepts in working memory were the primary processes accounting for suppression, peak suppression would be expected to be sustained until the sensory decision was made. Thus, although mechanisms involved in working memory and covert rehearsal cannot be completely ruled out, a number of findings are inconsistent with those explanations.

The present findings may be readily explained within the context of the proposed theoretical framework integrating speech perception theory, neurophysiological models, and theories of EEG generation (see Figure A-4). Direct-realists postulate that the motor system directly specifies the speech gestures that gave rise to the acoustic signal and constructivist theories propose that predictions about articulatory goals may function to constrain auditory perception (Callan et al., 2010). Within the context of these theories, results from this study clearly favor a constructivist account of sensorimotor integration and are consistent with dorsal stream, internal models of speech processing. More specifically, early beta suppression localized to the precentral gyrus prior to syllable discrimination, is consistent with an early forward model or neuronal population disinhibition that instantiates a general prediction about likely incoming sensory signals. As the subsequent percept is unknown, such an anticipatory model may be explained as neural tuning to expected acoustic features of the upcoming stimulus, similarly to the manner in which selective attention operates in visual perception (Hickok et al., 2011a).

Because early activity constitutes a prediction, continuous activity in the motor system to update the model would be expected until the initial set of articulatory hypotheses can be compared with online acoustic analysis. Peak activity just following acoustic analysis of the signal would be expected in the time period following acoustic input during which relevant features of the acoustic signal are matched with the initial forward constraints. Peak suppression would be expected only when the acoustic signal was sufficiently robust to be compared to the initial hypotheses. The finding that significant differences in  $\mu$  suppression between correct and chance trials occurred only immediately after stimulus offset may be explained by such a process. In other words, it appears that early motor models may be instantiated when discrimination is required but



fail to specify phonemes when auditory information is insufficient for comparison with the initial motor model.

However, given that internal articulatory models are thought to be developed via past experiences with producing speech (Callan et al., 2010; Hickok et al., 2011a; Skipper et al., 2006), early activity would not be expected for non-speech auditory signals as those signals have not been associated with vocal production. As such, suppression of the sensorimotor rhythm following tone processing may be, as suggested by others, an attempt to internally simulate pitch changes using cortical representations of the particulate vocal tract (Burton, 2009). It is worth noting that this notion would predict forward and inverse model pairs for not only speech sounds but for any sound repeatedly associated with goal-directed actions for producing the same sounds (Kohler et al., 2002). Consistent with this notion, a number of studies have demonstrated premotor and primary motor activation for sounds associated with motor sequences required to produce those sounds (Buchsbaum et al., 2011; Callan et al., 2006; D'Ausilio, Altenmüller, Olivetti Belardinelli, & Lotze, 2006; Dick, Lee, Nusbaum, & Price, 2011; Sammler, Harding, D'Ausilio, Fadiga, & Koelsch, 2010). As such, to further investigate the role of the motor system and perhaps internal models in general auditory perception, future investigations may employ learning tasks in which pitch change stimuli are paired with actions required to produce them.

As suggested in the first chapter, the present results favoring internal model concepts also fit well with those designed to explain event-related decreases in spectral power prior to a range of cognitive tasks. These early spectral changes have been associated with subsequent performance in memory, attention, and visual tracking. It has been suggested that alpha-like generators may reflect inhibitory filter mechanisms mediating top-down attentional control (Fellinger, Gruber, Zauner, Freunberger, & Klimesch, 2012; Fellinger, Klimesch, Gruber, Freunberger, & Doppelmayr, 2011; Hanslmayr, Gross, Klimesch, & Shapiro, 2011; Hari, 2006; Klimesch, 2011; Klimesch, Fellinger, & Freunberger, 2011). It is thought that power decreases (i.e., event-related decreases in power) signify a cortical release from inhibition that serves to ready the system for the coding of incoming information. This notion of release from inhibition has been extended to the function of the human MNS in sensorimotor integration (Pineda, 2005). According to theories of EEG generation (i.e., local/global theories), transient global coherence between multiple cortical generators may instantiate top-down anticipatory processes that facilitate subsequent sensory processing. These predictions may be further tested by exploring time-frequency changes in the other component clusters identified in the current study (see Appendix). Preliminary analysis indeed suggests wide spread entrainment across components emerging at near harmonics of thalamic rhythms (e.g., ~10Hz) with phase locked components across the traditional delta frequencies (0-5Hz). Analysis of other components of the EEG in this task may be critical to explaining how distributed neuronal systems share information for signal processing.

The present findings are also in accordance with the only studies that have investigated sensorimotor integration related to sublexical discrimination performance. A

previous combined fMRI and MEG study found that activity in precentral gyrus (vPMC) prior to and following stimulus onset was different for correct relative to incorrect syllable discrimination trials (Callan et al., 2010). Findings are also consistent with TMS studies showing that the PMC/M1 region modulates perceptual performance in auditory speech perception tasks, including acoustic segmentation and discrimination performance. Taken together, theory and experimental evidence strongly support the notion that sensorimotor integration via internal models plays an important role in the perception of speech.

### **Theoretical, Methodological, and Clinical Significance**

Neuroimaging evidence indicates that overlapping sensory and motor regions of the brain are involved in the perception and production of new percepts (Callan et al., 2003), monosyllables (Wilson et al., 2004), and in the receptive processing as well as generation of language (Pulvermüller & Fadiga, 2010). Evidence points to a dorsal sensorimotor stream specifying the manner in which phonemes are produced and a ventral stream involved in specifying the syntactic and semantic content of the incoming speech signal (Arbib, 2010; Hickok & Poeppel, 2004; 2007; Rauschecker & Scott, 2009; Skipper et al., 2006). During speech perception, the PMC and M1 regions of the dorsal stream have been shown to modulate perceptual performance in demanding speech processing tasks (D'Ausilio et al., 2009a; Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007; Sato, Tremblay, & Gracco, 2009). While these studies have provided critical information about cortical regions involved in perception, they have provided limited information about the time-course of sensorimotor processing relative to the onset of an acoustic event. As speech perception theory and current neurophysiological models predict differential sensorimotor involvement relative to the onset of an acoustic stimulus, higher temporal resolution is critical to determining how sensorimotor integration functions in speech processing.

The current study is the first to employ a BSS method used previously in EEG studies of visual perception to address many of the limitations associated with traditional studies of acoustic speech processing. Those limitations include insufficient spatiotemporal resolution, lack of evidence for a functional role, and lack of specificity to speech as opposed to general auditory processing. The present study favors dynamic, internal model concepts of speech processing over computational mechanisms proposed by direct-realists. In addition, the study provides the first evidence supporting claims that these internal models operate similarly to a kind of phonological or articulatory selective attention (Callan et al., 2010; Hickok et al., 2011; Skipper et al., 2006). For example, the finding that both correct and chance syllable discrimination trials were preceded by early  $\mu$  suppression is what would be expected if forward articulatory models function similarly to selective attention (Callan et al., 2010; Hickok et al., 2011a). That is, if early articulatory hypotheses function in a manner similar to attention, early motor activity would be expected regardless of subsequent correct or chance level performance.

The study also provides the first evidence that early forward models are related to perceptual performance at the point in time when acoustic features are sufficient for comparison with initial hypotheses (i.e., immediately following acoustic stimuli). Furthermore, this study suggests that internal models are specific to a syllable discrimination task relative to a similar, rapid pitch change task. As both tasks required attention to a rapid transitional acoustic cue for successful discrimination, general attentional mechanisms cannot account for differences in early  $\mu$  rhythm suppression for speech as opposed to tone stimuli. This finding is critical to determining underlying mechanisms, as no studies have measured the ongoing time-course of motor activity for speech and non-speech control tasks with similar discrimination requirements. As such, results favor articulatory mechanisms that are either specific to speech (see Liberman & Whalen, 2000 for discussion) or alternatively to any auditory signal previously paired with actions (Kholer et al., 2002).

However, because specificity of sensory and motor processing to speech has traditionally been another area of heated debate, this issue deserves more explicit comment. One particular view of speech processing, called the ‘motor theory’ of speech perception postulated the use of articulatory goals to mediate and constrain perception long ago (Callan et al., 2010). According to Liberman’s ‘motor theory’ the lack of invariance problem is solved via a specialized phonetic module evolved to track intended invariant articulatory targets as opposed to acoustic features. The motor theory predicts that this articulatory phonetic module is critical for speech perception and thus motor activity should be present in all contexts. However, several lines of evidence are inconsistent with a specialized process that is both necessary and sufficient for perceptual invariance, including categorical perception in non-human species (Kuhl, 1976), lesion evidence (Lotto et al., 2009) and neurophysiological evidence suggesting that speech processing is not always accompanied by activity in motor regions (Skipper et al., 2005; Skipper et al., 2006). The lack of strong motor activity in passive speech perception in this study is consistent with some previous neuroimaging studies and is somewhat contrary to the predictions of the motor theory. As such, while the present study indicates sensorimotor activity specific to a syllable discrimination task, it also favors a more dynamic concept of speech perception in which sensorimotor integration via a transiently interactive neuronal system may function to aid acoustic analysis depending on context.

This dynamic theory of speech processing (DTS) may be derived from concepts of perception in which cognitive mechanisms critical for interspecies communication developed in noisy, real-world environments (Gallantucci et al., 2006; Wilson, 2009). As such, communication and the cognitive mechanisms that support it maintain strong ties to the environmental context in which it developed. The concept of parity (i.e., ‘what counts as communication for you also counts for me’) may be seen as an attempt to integrate the contexts in which cognition for the purposes of within species communication must be understood (Galantucci et al., 2006). First, as parity is intended to constrain speaker and listener interactions in ecological contexts, it allows for agreement between the sender and receiver of information. Second, because the living systems in which parity evolved are more likely to be achieved if perceptual and motor systems co-evolve, it makes sense that cognition retains these relationships. These two concepts highlight the connection

between the evolution of sensorimotor systems for online action and those that evolved or developed for offline sensory analysis. Broadly, the DTS may be discussed under the umbrella of a controversial concept known as embodied cognition (EC). Much like the concept of parity in communication, EC views the evolution and development of cognitive processes as inextricable from the body that implements them.

Because speaking environments are often noisy and involve multiple speakers and listeners, it is not surprising that the brain has exapted more than one mechanism for the parsing of speech and other communicative signals (Crawcour et al., 2009; D'Ausilio et al., 2010; Skipper et al., 2007; Wilson et al., 2009). As such, DTS and other active theories of speech processing, predict that during real-world communicative interactions, the relative weight of sensory and motor subsystems shift transiently depending on the information available and expectations between communication partners (Crawcour et al., 2009; Skipper et al., 2006; van Wassenhove et al., 2005). A DTS might provide a framework for explaining a number of outstanding questions in the neuroscience of speech processing. It might explain why motor and sensory systems are differentially active in some speech processing tasks while conspicuously absent in others. Further, such a framework might provide an explanation for why motor systems are active when sensory processing alone may be sufficient. For example, dual-stream theories (Arbib, 2010; Hickok & Poeppel, 2007) specify that the ventral stream only should be active when participants perform tasks in which syntactic, semantic, and phonological features are critical for performance. The dorsal stream should be active only when features of the signal critical for task performance are gestural, visual, or acoustic (Hickok & Poeppel, 2007; Skipper et al., 2006). That is, regional brain activation patterns are associated with task specific findings in which active attention to features of the communicative signal are critical for task performance. These same patterns active in some contexts are often absent when the task requires attention to a different set of features (e.g., acoustic vs. linguistic tasks) (Hickok & Poeppel, 2007). This phenomenon is known as a double-disassociation between linguistic and purely perceptual speech processing tasks (Hickok & Poeppel, 2004).

According to DTS, the lack of invariance problem may be solved or effectively avoided via active attention to features of the incoming sensory information that are critical for successful interpretation. This model of speech processing predicts that during face-to-face communicative interactions and particular circumstances in which bottom-up information is insufficient to resolve the speech signal (e.g., noisy conditions) transient attention to critical features instantiates top-down constraints on perception. These critical features include a wide range of observable gestures known to influence perception in experimental contexts (e.g., head, jaw, lip and facial movements) and experience dependent syntactic, semantic, and phonological knowledge (Skipper et al., 2006). Critical features may also include value based judgments involving regions of the limbic system and autonomic nervous system thought to be important for identifying prosodic features, speaker characteristics, and pragmatics. Thus, as opposed to direct matching via a phonetic module or direct matching between acoustic features and articulatory gestures, a distributed neuronal system utilizes experiential knowledge from language and speech production to decrease the inherent ambiguity of the communicative

signal. Although this concept does not specify an exact computational mechanism for solving the ‘lack of invariance problem’, it provides predictions that are testable via dynamic computational models and the current EEG methodology.

## **Future Directions**

In real-world contexts, a DTS predicts that multisensory and motor features of the speech signal are likely to be transiently involved in the receptive processing of speech and ultimately language comprehension. A complementary unifying theory of neural processing called dynamic field theory (DFT) might provide a framework for the prediction and interpretation of the behavior of neural populations in DTS. The DFT models the neural interaction in dynamic multilayer fields that account for the properties of sensory and motor systems in working memory, attention, and long-term memory. The DFT predicts that cognition is both embodied in systems evolved for sensing, acting, and in environmental context. As such, it predicts a relationship of the kind described both by Hickok et al. (2011) and Callan et al. (2010) in which a kind of phonological selective attention modulates activity in sensory areas. For example, a similar function for the prefrontal cortex has been proposed for visual movement processing and the behavior of such a system can be modeled using multilayer DFT (Spencer, Perone, & Johnson, 2010). Moreover, the DFT might explain findings of motor activity under conditions in which segments of the acoustic signal must be held in working memory (Burton et al., 2009; Sato et al., 2009) and in which new percepts (e.g., non-native contrasts) must be compared with long-term motor-phonological representations (Callan et al., 2003). Further, as the DFT predicts relationships between action systems and cognition generally, it can incorporate findings of motor activity to acoustic processing related to reproducible actions (Rauschecker & Scott, 2009). Finally, as the DFT was formulated to predict the activity of neuronal populations and EEG is thought to be a measure of local/global population dynamics, the EEG may be an effective way to provide the computational model with supporting data.

As transient attention to critical features is predicted to be important for extracting communicative intent, high-temporal resolution, ongoing measures are necessary to test the predictions of DTS. Future investigations utilizing the current methodology may make use of multiple psychophysiological measures in both controlled experimental and real-world contexts to investigate the role of active processing. These approaches will include combined measures of infrared eye-tracking, measures of autonomic arousal, and EEG. Eye-gaze is thought to be a measure of focused attention as the quality of visual resolution drops off rapidly outside the center of pupillary focus (Henderson, 2003). Saccadic eye-movements signaling changes in attention occur on a millisecond time scale. As such, changes in attention signaled by gaze focus may be used as event-structures around which EEG can be epoched, allowing for an examination of neuronal activity surrounding the attentional event. As these transient shifts are thought to be important for linguistic comprehension in real-world contexts, this experimental paradigm may also be used to examine changes in attention thought to influence perception, including active processing of head movements, manual gestures, tongue, mandibular,



and lip movements. Finally, because value based judgments are also thought to be important to speech and language processing, neuronal population dynamics localized to limbic regions may be measured simultaneously with autonomic measures of arousal known to index changes in emotional state (Bowers, Saltuklaroglu, & Kalinowski, 2011). The auditory dorsal stream correlates of transient shifts in attention to critical features may be also be examined in psychoacoustic phenomena in which the motor system is thought to play a critical role in the behavioral outcome such as the McGurk-McDonald effect (McGurk & MacDonald, 1976), trading relations (Repp, 1982), duplex perception (Repp, 1984), and the identification of sine-wave speech. Sine-wave speech is generated by replacing the formants critical for speech recognition with sine-waves (Benson, Richardson, Whalen, & Lai, 2006) and is of particular interest in that whether or not participants anticipate speech significantly shapes whether or not sine-wave analogs are recognizable as speech.

### **Possible Clinical Implications**

As processes of expectation or attention are thought to be important for perceiving speech, it may be important to examine whether violation of expectations for communicative interactions more heavily rely on the proposed internal model mechanisms. This possibility might be examined using disordered speech such as dysarthria, stuttering, or apraxia. The computational mechanisms by which communication partners (CP) perceive stuttering are particularly interesting as CP reactions to stuttered speech may significantly influence the development of covert features of the pathology (Bowers, Crawcour, Saltuklaroglu, & Kalinowski, 2009; Bowers et al., 2011). People who stutter often avoid certain people, places, and social events due the expectation of negative CP reactions. Equally as interesting is the possibility that the current methodology might be used shed light on the computational mechanisms underlying the fluency enhancing effects of various auditory signals in people who stutter (PWS). It has been proposed that disrupted internal models during speech production in PWS might be reinstated by auditory signals specifying intended articulatory gestures via the auditory dorsal stream (Hickok et al., 2011; but see Brown, Ingham, Ingham, Laird, & Fox, 2005; Max et al., 2004; Saltuklaroglu & Kalinowski, 2006 for similar accounts).

As early changes in spectral power are thought to influence subsequent perceptual performance in noisy conditions, it is likely that some populations with difficulty resolving percepts in noise might utilize sensorimotor integration in a compensatory capacity. Older adult populations without hearing loss have been shown to recruit areas of the motor system and regions thought to be involved in attention more heavily than younger listeners (Wong et al., 2009). The motor system may function similarly to aid perception in children who have deficits in the perception of speech or spectotemporal analysis generally but have intact speech production (Benasich, Thomas, Choudhury, & Leppänen, 2002). It is an open question whether the articulatory system plays an important role in perception for populations with hearing impairment. For instance, do children with hearing impairment who develop spared speech production utilize internal

constrains on perception to a greater extent than those who do not have hearing impairment? Further as deficits or differences in sensorimotor integration may play an important role in specific language impairment and autism spectrum disorders (LeBel, Sharma, & Pineda, 2009), the current methodology may provide information important for treatment. Establishing the time-frequency features of the EEG for both normal and clinical populations may be critical, as these signals have been shown to change as a function of experience and may thus be used as neuromodulatory feedback to enhance current therapeutic protocols (Cannon, Lubar & Baldwin, 2008; Congedo et al., 2008; Pineda, 2005; Van Der Loo, Congedo, Plazier, Van De Heyning & De Ridder, 2007).

## **Limitations**

Although these findings present compelling evidence that the  $\mu$  rhythm has an important function in the perception of speech, several limitations must be considered. First, as this study used only meaningless syllable and tone stimuli in the auditory modality only, the current findings are limited in the extent to which they can be generalized to other contexts. Second, as only 32 channels were used in this study, it may be important to investigate whether improved spatial resolution may be obtained for larger electrode arrays (i.e., 64-128 channels). However, informal comparisons between smaller and larger arrays have generally shown similar independent components associated with the most synchronous activity of the EEG (Onton & Makeig, 2006) and strong correlations between dipole models for low density arrays and BOLD activity have been demonstrated (Debener, Ullsperger, Siegel, & Engel, 2006). Second, given that this study did not employ a digitizer to input the exact location of electrodes (as opposed to relying on the average 10-20 method) and employed relatively imprecise head models, it may be important to replicate the current findings using more precise limiting parameters for source localization.

Third, although findings from this study implicate the sensorimotor rhythm in accurate speech perception, performance was manipulated at the extremes so that performance was either accurate or not better than chance. Future studies might demonstrate that the sensorimotor rhythm tracks perceptual discrimination performance across multiple SNRs. Callan et al., (2010) found that correct trials were preceded and followed by significant low and high frequency ERD (i.e., suppression) emerging from the vPMC, suggesting that this region plays an important role in processing speech percepts when bottom-up, sensory driven processes fail. However, another possible explanation offered for that result is that correct trials are simply associated with motor activity whereas incorrect trials are not due to the fact that precepts were phenomenologically ambiguous. As attention (i.e., active processing) is thought to play a critical role in activating the dorsal stream network (see Skipper et al., 2006), it is possible that participants were simply not attending during incorrect trials. In other words, correct trials are simply accompanied by early and late activation in the precentral gyrus as opposed to functional involvement in correct discrimination.



Findings in the present study may help to flesh out the underlying role of early forward models. The presence of early beta suppression prior to stimulus onset and throughout stimulus processing would suggest that the motor system was involved at all stages of processing for speech stimuli, both for conditions in which percepts were discriminable and for conditions in which they were not. Thus, it appears that forward models may participate in perception even when the acoustic signal is so impoverished that discrimination between phonemes is not possible. Further, it appears that this process is specific to incoming acoustic signals that have been paired previously with articulatory movements. In light of the current findings, it seems possible that early forward models may function similarly to effort in attention. That is, when stimuli are attended regardless of the likelihood of success or failure, early motoric models are instantiated followed by an attempt at synthesis. It is likely that synthesis or sensorimotor integration is achieved only when early articulatory hypotheses and subsequent acoustic cues are sufficient to specify phonemic units.

## LIST OF REFERENCES

- Altschuler E.L., Vankov, A., Wang, V. Ramachandran V. and Pineda, J.A. Person see, person do cells, *Abstracts-Social Neuroscience* 23 (1997), p. 1848.
- Arbib, M. A. (2010). Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain and Language*, 112(1), 12–24. doi: 10.1016/j.bandl.2009.10.001.
- Aziz-Zadeh, L., & Ivry, R. B. (2009). The human mirror neuron system and embodied representations. *Advances in Experimental Medicine and Biology*, 629, 355–376. doi:10.1007/978-0-387-77064-2\_18.
- Baar-Eroglu, C., Baar, E., Demiralp, T., & Schürmann, M. (1992). P300-response: possible psychophysiological correlates in delta and theta frequency channels. A review. *International Journal of Psychophysiology*, 13(2), 161–179.
- Basar, E., Basar-Eroglu, C., Karakas, S., & Schurmann, M. (2001a). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, 39(2-3), 241–248.
- Basar, E., Basar-Eroglu, C., Karakas, S., & Schurmann, M. (2001b). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, 39(2-3), 241–248.
- Basar, E., Schurmann, M., Basar-Eroglu, C., & Karakas, S. (1997). Alpha oscillations in brain functioning: an integrative theory. *International Journal of Psychophysiology*, 26(1-3), 5–29.
- Basar-Eroglu, C., Struber, D., Schurmann, M., Stadler, M., & Basar, E. (1996). Gamma-band responses in the brain: a short review of psychophysiological correlates and functional significance. *International Journal of Psychophysiology*, 24(1-2), 101–112.
- Benasich, A. A., Thomas, J. J., Choudhury, N., & Leppänen, P. H. . (2002). The importance of rapid auditory processing abilities to early language development: evidence from converging methodologies. *Developmental psychobiology*, 40(3), 278–292.
- Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, 25(1), 60–83.
- Benson, R. R., Richardson, M., Whalen, D. H., & Lai, S. (2006). Phonetic processing areas revealed by sinewave speech and acoustically similar non-speech. *Neuroimage*, 31(1), 342–353.

- 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. *Nature Neuroscience*, 7(3), 295–301.
- Binder, J. R., Rao, S. M., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Bandettini, P. A., Wong, E. C., et al. (1994). Functional magnetic resonance imaging of human auditory cortex. *Annals of Neurology*, 35(6), 662–672.
- Bowers, A., Crawcour, S., Saltuklaroglu, T., & Kalinowski, J. (2009). Gaze aversion to stuttered speech: a pilot study investigating differential visual attention to stuttered and fluent speech. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, 1. doi:10.1080/13682820902763951.
- Bowers, A., Saltuklaroglu, T., & Kalinowski, J. (2011). Autonomic arousal in adults who stutter prior to various reading tasks intended to elicit changes in stuttering frequency. *International Journal of Psychophysiology*.
- Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human brain mapping*, 25(1), 105–117.
- Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language*, 89(2), 370–376. doi:10.1016/S0093-934X(03)00356-0.
- Buccino, G., Solodkin, A., & Small, S. L. (2006). Functions of the mirror neuron system: implications for neurorehabilitation. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology*, 19(1), 55–63.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - An aggregate analysis of lesion and fMRI data. *Brain and Language*. doi:10.1016/j.bandl.2010.12.001.
- Burton, M. W. (2009). Understanding the role of the prefrontal cortex in phonological processing. *Clinical Linguistics & Phonetics*, 23(3), 180–195. doi:10.1080/02699200802394963.
- Burton, M. W., LoCasto, P. C., Krebs-Noble, D., & Gullapalli, R. P. (2005). A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *NeuroImage*, 26(3), 647–661. doi:10.1016/j.neuroimage.2005.02.024.
- Caggiano, V., Fogassi, L., Rizzolatti, G., Thier, P., & Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science (New York, N.Y.)*, 324(5925), 403–406. doi:10.1126/science.1166818.

- 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, and Hearing Research: JSLHR*, *43*(3), 721–736.
- Callan, D.E., Callan, A., Gamez, M., Sato, M., & Kawato, M. (2010). Premotor cortex mediates perceptual performance. *NeuroImage, In Press, Corrected Proof*. doi:10.1016/j.neuroimage.2010.02.027.
- Callan, D. E., Callan, A. M., Kroos, C., & Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. *Cognitive Brain Research*, *10*(3), 349–353.
- 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:10.1016/j.neuroimage.2006.01.036.
- Cannon, R., Lubar, J., & Baldwin, D. (2008). Self-perception and experiential schemata in the addicted brain. *Applied Psychophysiology and Biofeedback*, *33*(4), 223–238. doi:10.1007/s10484-008-9067-9.
- Cattaneo, L., & Rizzolatti, G. (2009a). The Mirror Neuron System. *Arch Neurol*, *66*(5), 557–560. doi:10.1001/archneurol.2009.41.
- Cattaneo, L., & Rizzolatti, G. (2009b). The mirror neuron system. *Archives of Neurology*, *66*(5), 557–560. doi:10.1001/archneurol.2009.41.
- 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:10.1016/j.clinph.2008.09.007.
- Conway, B. A., Halliday, D. M., Farmer, S. F., Shahani, U., Maas, P., Weir, A. I., & Rosenberg, J. R. (1995). Synchronization between motor cortex and spinal motoneuronal pool during the performance of a maintained motor task in man. *The Journal of physiology*, *489*(Pt 3), 917–924.
- Corballis, M. C. (2010). Mirror neurons and the evolution of language. *Brain and Language*, *112*(1), 25–35. doi:10.1016/j.bandl.2009.02.002.

- 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:10.1016/j.neuropsychologia.2009.05.001.
- Crone, N. E., Boatman, D., Gordon, B., & Hao, L. (2001). Induced electrocorticographic gamma activity during auditory perception. *Clinical Neurophysiology*, *112*(4), 565–582. doi:10.1016/S1388-2457(00)00545-9.
- D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, *24*(3), 955–958.
- D'Ausilio, A., Bufalari, I., Salmas, P., Busan, P., & Fadiga, L. (2011). Vocal pitch discrimination in the motor system. *Brain and Language*.
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009a). The motor somatotopy of speech perception. *Current Biology: CB*, *19*(5), 381–385. doi:10.1016/j.cub.2009.01.017.
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009b). The motor somatotopy of speech perception. *Current Biology: CB*, *19*(5), 381–385. doi:10.1016/j.cub.2009.01.017.
- 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:10.1016/j.heares.2007.01.014.
- Debener, S., Ullsperger, M., Siegel, M., & Engel, A. K. (2006). Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends in cognitive sciences*, *10*(12), 558–563.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9–21.
- Dick, A. S., Goldin-Meadow, S., Hasson, U., Skipper, J. I., & Small, S. L. (2009). Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping*, *30*(11), 3509–3526. doi:10.1002/hbm.20774.
- Dick, F., Lee, H. L., Nusbaum, H., & Price, C. J. (2011). Auditory-motor expertise alters “speech selectivity” in professional musicians and actors. *Cerebral Cortex*, *21*(4), 938–948.
- Elul, R. (1972). The genesis of the EEG. *Int. Rev. Neurobiol*, *15*(2), 227.

- Ernst, S. M. A., Verhey, J. L., & Uppenkamp, S. (2008). Spatial dissociation of changes of level and signal-to-noise ratio in auditory cortex for tones in noise. *NeuroImage*, *43*(2), 321–328. doi:10.1016/j.neuroimage.2008.07.046.
- Fellinger, R., Klimesch, W., Gruber, W., Freunberger, R., & Doppelmayr, M. (2011). Pre-stimulus alpha phase-alignment predicts P1-amplitude. *Brain Research Bulletin*, *85*(6), 417–423. doi:10.1016/j.brainresbull.2011.03.025.
- Fellinger, R., Gruber, W., Zauner, A., Freunberger, R., & Klimesch, W. (2012). Evoked traveling alpha waves predict visual-semantic categorization-speed. *NeuroImage*, *59*(4), 3379–3388. doi:10.1016/j.neuroimage.2011.11.010.
- Fisch, B. J. (1999). *Fisch and Spehlmann's EEG primer: basic principles of digital and analog EEG*. Elsevier Amsterdam.
- Formaggio, E., Storti, S. F., Avesani, M., Cerini, R., Milanese, F., Gasparini, A., Acler, M., et al. (2008). EEG and fMRI coregistration to investigate the cortical oscillatory activities during finger movement. *Brain Topography*, *21*(2), 100–111. doi:10.1007/s10548-008-0058-1.
- Fowler, C. (n.d.). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, (14), 3–28.
- Fowler, C. A. (1996). Listeners do hear sounds» not tongues. *J. Acoust. Soc. Am*, *99*, 3.
- Fowler, C. A., Brown, J. M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks *Journal of Memory and Language*, *49*(3), 396–413.
- Fridriksson, J., Moser, D., Ryalls, J., Bonilha, L., Rorden, C., & Baylis, G. (2009). Modulation of frontal lobe speech areas associated with the production and perception of speech movements. *Journal of Speech, Language, and Hearing Research: JSLHR*, *52*(3), 812–819. doi:10.1044/1092-4388(2008/06-0197).
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*(3), 361–377.
- Gallese, V., Gernsbacher, M. A., Heyes, C., Hickok, G., & Iacoboni, M. (2011). Mirror neuron forum. *Perspectives on Psychological Science*, *6*(4), 369–407.
- Garbarini, F., & Adenzato, M. (2004). At the root of embodied cognition: Cognitive science meets neurophysiology. *Brain and Cognition*, *56*(1), 100–106.
- Gibson, J. J. (1986). *The ecological approach to visual perception*.
- Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. ., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, *56*(6), 1127–1134.



- Goncalves, S. I., De Munck, J. C., Pouwels, P. J. W., Schoonhoven, R., Kuijter, J. P. A., Maurits, N. M., Hoogduin, J. M., et al. (2006). Correlating the alpha rhythm to BOLD using simultaneous EEG/fMRI: inter-subject variability. *Neuroimage*, *30*(1), 203–213.
- Graimann, B., & Pfurtscheller, G. (2006). Quantification and visualization of event-related changes in oscillatory brain activity in the time-frequency domain. *Progress in Brain Research*, *159*, 79–97. doi:10.1016/S0079-6123(06)59006-5.
- Grin-Yatsenko, V. A., Baas, I., Ponomarev, V. A., & Kropotov, J. D. (2010). Independent component approach to the analysis of EEG recordings at early stages of depressive disorders. *Clinical Neurophysiology*, *121*(3), 281–289. doi:10.1016/j.clinph.2009.11.015.
- Guenther, F. H. (1995). A modeling framework for speech motor development and kinematic articulator control. *Proceedings of the XIIIth International Conference of Phonetic Sciences* (Vol. 2, pp. 92–99).
- Guenther, F. H. (2001). Neural modeling of speech production. *Proceedings of the 4th international Nijmegen speech motor conference* (pp. 13–16).
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, *39*(5), 350–365.
- Handel, S. (1989). Listening: An introduction to the perception of auditory events.
- Hanslmayr, S., Gross, J., Klimesch, W., & Shapiro, K. L. (2011). The role of alpha oscillations in temporal attention. *Brain Research Reviews*, *67*(1-2), 331–343. doi:10.1016/j.brainresrev.2011.04.002.
- Hari, R., Salmelin, R., Mäkelä, J. P., Salenius, S., & Helle, M. (1997). Magnetoencephalographic cortical rhythms. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *26*(1-3), 51–62.
- Hari, R., (2006). Action-perception connection and the cortical mu rhythm. *Event-Related Dynamics of Brain Oscillations* (Vol. Volume 159, pp. 253–260). Elsevier. Retrieved from <http://www.sciencedirect.com/science/article/B7CV6-4M6SMB6-N/2/52b6fbe4899453d9c40676d29a396bd1>.
- Hari, R., Levänen, S., & Raij, T. (2000). Timing of human cortical functions during cognition: role of MEG. *Trends in Cognitive Sciences*, *4*(12), 455–462. doi:10.1016/S1364-6613(00)01549-7.
- Hasson, U., Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2007). Abstract coding of audiovisual speech: beyond sensory representation. *Neuron*, *56*(6), 1116–1126. doi:10.1016/j.neuron.2007.09.037.



- Hatfield, G. (2002). Psychology, philosophy, and cognitive science: Reflections on the history and philosophy of experimental psychology. *Mind & language*, 17(3), 207–232.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in cognitive sciences*, 7(11), 498–504.
- Hickok, G., Houde, J., & Rong, F. (2011a). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron*, 69(3), 407–422.
- Hickok, G., Houde, J., & Rong, F. (2011b). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron*, 69(3), 407–422.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4(4), 131–138.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67–99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.
- Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews*, 6(3), 121–143. doi:10.1016/j.plrev.2009.06.001.
- Hickok, G., Holt, L. L., & Lotto, A. J. (2009). Response to Wilson: What does motor cortex contribute to speech perception? *Trends in Cognitive Sciences*, 13(8), 330–331. doi:10.1016/j.tics.2009.05.002.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, 8(5), 393–402. doi:10.1038/nrn2113.
- Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews. Neuroscience*, 7(12), 942–951. doi:10.1038/nrn2024.
- Jasper, H., & Penfield, W. (1949). Electrocorticograms in man: effect of voluntary movement upon the electrical activity of the precentral gyrus. *European Archives of Psychiatry and Clinical Neuroscience*, 183(1), 163–174.
- 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.

- Joanisse, M. F., & Gati, J. S. (2003). Overlapping neural regions for processing rapid temporal cues in speech and nonspeech signals. *NeuroImage*, *19*(1), 64–79.
- 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.
- Kalinowski, J., & Saltuklaroglu, T. (2003). Speaking with a mirror: engagement of mirror neurons via choral speech and its derivatives induces stuttering inhibition. *Medical Hypotheses*, *60*(4), 538–543.
- Karrasch, M., Krause, C. M., Laine, M., Lang, A. H., & Lehto, M. (1998). Event-related desynchronization and synchronization during an auditory lexical matching task. *Electroencephalography and Clinical Neurophysiology*, *107*(2), 112–121.
- Kent, R. D. (1997). *The speech sciences*. Singular Pub. Group.
- Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters*, *244*(2), 73–76. doi:10.1016/S0304-3940(98)00122-0.
- Klimesch, W. (2011). Evoked alpha and early access to the knowledge system: the P1 inhibition timing hypothesis. *Brain Research*, *1408*, 52–71. doi:10.1016/j.brainres.2011.06.003.
- Klimesch, W., Fellinger, R., & Freunberger, R. (2011). Alpha oscillations and early stages of visual encoding. *Frontiers in Psychology*, *2*, 118. doi:10.3389/fpsyg.2011.00118.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, *297*(5582), 846–848. doi:10.1126/science.1070311.
- Krause, C.M., Sillanmäki, L., Koivisto, M., Saarela, C., Häggqvist, A., Laine, M., & Hämäläinen, H. (2000). The effects of memory load on event-related EEG desynchronization and synchronization. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *111*(11), 2071–2078.
- Krause, C. M., (2006). Cognition- and memory-related ERD/ERS responses in the auditory stimulus modality. *Progress in Brain Research*, *159*, 197–207. doi:10.1016/S0079-6123(06)59013-2.
- Kuhl, P. K. (1976). Speech perception by the chinchilla: Categorical perception of synthetic alveolar plosive consonants. *The Journal of the Acoustical Society of America*, *60*, S81.

- 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.
- Le Bel, R. M., Pineda, J. A., & Sharma, A. (2009). Motor-auditory-visual integration: The role of the human mirror neuron system in communication and communication disorders. *Journal of Communication Disorders*, *42*(4), 299–304. doi:10.1016/j.jcomdis.2009.03.011.
- Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural computation*, *11*(2), 417–441.
- Liberman, A. M. (1957). Some Results of Research on Speech Perception. *The Journal of the Acoustical Society of America*, *29*(1), 117–123. doi:10.1121/1.1908635.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1–36. doi:10.1016/0010-0277(85)90021-6.
- Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*(5), 187–196. doi:10.1016/S1364-6613(00)01471-6.
- LoCasto, P. C., Krebs-Noble, D., Gullapalli, R. P., & Burton, M. W. (2004a). An fMRI investigation of speech and tone segmentation. *Journal of Cognitive Neuroscience*, *16*(9), 1612–1624. doi:10.1162/0898929042568433.
- LoCasto, P. C., Krebs-Noble, D., Gullapalli, R. P., & Burton, M. W. (2004b). An fMRI investigation of speech and tone segmentation. *Journal of Cognitive Neuroscience*, *16*(9), 1612–1624. doi:10.1162/0898929042568433.
- Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, *13*(3), 110–114. doi:10.1016/j.tics.2008.11.008.
- Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and Clinical Neurophysiology*, *86*(4), 283–293.
- Makeig, S., Bell, Tzyy-ping Jung, & Terrence J. Sejnowski. (1996). Independent Component Analysis of Electroencephalographic Data. MIT Press. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=?doi=10.1.1.36.4803>.
- Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210.

- Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(20), 10979–10984.
- Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002a). Dynamic brain sources of visual evoked responses. *Science*, *295*(5555), 690.
- Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002b). Dynamic Brain Sources of Visual Evoked Responses. *Science*, *295*(5555), 690–694. doi:10.1126/science.1066168.
- Makeig, S., Westerfield, M., Townsend, J., Jung, T. P., Courchesne, E., & Sejnowski, T. J. (1999). Functionally independent components of early event-related potentials in a visual spatial attention task. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *354*(1387), 1135–1144. doi:10.1098/rstb.1999.0469.
- Manshanden, I., De Munck, J. C., Simon, N. R., & Lopes da Silva, F. H. (2002). Source localization of MEG sleep spindles and the relation to sources of alpha band rhythms. *Clinical neurophysiology*, *113*(12), 1937–1947.
- Marco-Pallarés, J., Grau, C., & Ruffini, G. (2005). Combined ICA-LORETA analysis of mismatch negativity. *NeuroImage*, *25*(2), 471–477. doi:10.1016/j.neuroimage.2004.11.028.
- Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. The MIT Press. Retrieved from [http://books.google.com/books?hl=en&lr=&id=ksYnXu6HnU0C&oi=fnd&pg=PR5&dq=massaro,+1998&ots=QSB3pCUEWn&sig=8hctyKyggzkDv9Nq6eeUE\\_g2M\\_s](http://books.google.com/books?hl=en&lr=&id=ksYnXu6HnU0C&oi=fnd&pg=PR5&dq=massaro,+1998&ots=QSB3pCUEWn&sig=8hctyKyggzkDv9Nq6eeUE_g2M_s).
- Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=?doi=10.1.1.4.3841>.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices.
- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, *17*(19), 1692–1696.
- Meister, Ingo G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology: CB*, *17*(19), 1692–1696. doi:10.1016/j.cub.2007.08.064.

- Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, *14*(17), 2135.
- Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2010). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*.
- Muthukumaraswamy, S. D., & Johnson, B. W. (2004). Changes in rolandic mu rhythm during observation of a precision grip. *Psychophysiology*, *41*(1), 152–156.
- Muthukumaraswamy, S. D., & Johnson, B. W. (2004). Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *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 Research*, *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. *Cognitive Brain Research*, *19*(2), 195–201. doi:10.1016/j.cogbrainres.2003.12.001.
- Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(2), 913–918.
- Nunez, P. L. (2000). Toward a quantitative description of large-scale neocortical dynamic function and EEG. *Behavioral and Brain Sciences*, *23*(03), 371–398.
- Nunez, P. L., & Silberstein, R. B. (2000). On the relationship of synaptic activity to macroscopic measurements: Does co-registration of EEG with fMRI make sense? *Brain Topography*, *13*(2), 79–96.
- Nunez, P. L., & Srinivasan, R. (2006). A theoretical basis for standing and traveling brain waves measured with human EEG with implications for an integrated consciousness. *Clinical neurophysiology*, *117*(11), 2424–2435.
- 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 Research. Cognitive Brain Research*, *24*(2), 190–198. doi:10.1016/j.cogbrainres.2005.01.014.
- Oberman, L. M., Pineda, J. A., & Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Social Cognitive and Affective Neuroscience*, *2*(1), 62–66. doi:10.1093/scan/nsi022.

- 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.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.
- Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. *Event-Related Dynamics of Brain Oscillations* (Vol. Volume 159, pp. 99–120). Elsevier. Retrieved from <http://www.sciencedirect.com/science/article/B7CV6-4M6SMB6-9/2/9717aa75e6e60e6c6d091af2d7a4b2e8>.
- 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. *Human brain mapping*, *17*(3), 179–192.
- Osipova, D., Hermes, D., Jensen, O., & Rustichini, A. (2008). Gamma power is phase-locked to posterior alpha activity. *PLoS One*, *3*(12), e3990.
- Osnes, B., Hugdahl, K., Hjelmervik, H., & Specht, K. (2012). Stimulus expectancy modulates inferior frontal gyrus and premotor cortex activity in auditory perception. *Brain and Language*.
- Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, *24 Suppl D*, 5–12.
- Pfurtscheller, G. (1999). Quantification of ERD and ERS in the time domain. *Event-related desynchronization*, *6*, 89–105.
- Pfurtscheller, G., & Klimesch, W. (1990). Topographical display and interpretation of event-related desynchronization during a visual-verbal task. *Brain Topography*, *3*(1), 85–93.
- Pfurtscheller, G., Stancák, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band--an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *24*(1-2), 39–46.
- Pineda, J. A. (2005). The functional significance of mu rhythms: translating “seeing” and “hearing” into “doing.” *Brain Research. Brain Research Reviews*, *50*(1), 57–68. doi:10.1016/j.brainresrev.2005.04.005.
- Pizzagalli, D. A., Oakes, T. R., & Davidson, R. J. (2003). Coupling of theta activity and glucose metabolism in the human rostral anterior cingulate cortex: An EEG/PET study of normal and depressed subjects. *Psychophysiology*, *40*(6), 939–949.



- Price, C. J. (2010). ANNALS OF THE NEW YORK ACADEMY OF SCIENCES Issue: The Year in Cognitive Neuroscience The anatomy of language: a review of 100 fMRI studies.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci*, *11*(5), 351–360. doi:10.1038/nrn2811.
- 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.
- Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, *92*(1), 81.
- Repp, B. H. (1984). Against a role of “chirp” identification in duplex perception. *Attention, Perception, & Psychophysics*, *35*(1), 89–93.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194.
- Salenius, S., Kajola, M., Thompson, W. L., Kosslyn, S., & Hari, R. (1995). Reactivity of magnetic parieto-occipital alpha rhythm during visual imagery. *Electroencephalography and Clinical Neurophysiology*, *95*(6), 453–462.
- Salenius, S., Salmelin, R., Neuper, C., Pfurtscheller, G., & Hari, R. (1996). Human cortical 40 Hz rhythm is closely related to EMG rhythmicity. *Neuroscience Letters*, *213*(2), 75–78.
- Sammler, D., Harding, E. E., D’Ausilio, A., Fadiga, L., & Koelsch, S. (2010). Music and action: do they share neural resources? *11th International Conference of Music Perception and Cognition*.
- Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T., & Simola, J. (1991). Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience Letters*, *127*(1), 141–145.
- Sato, J. R., Rondinoni, C., Sturzbecher, M., de Araujo, D. B., & Amaro, E. (2010). From EEG to BOLD: brain mapping and estimating transfer functions in simultaneous EEG-fMRI acquisitions. *NeuroImage*, *50*(4), 1416–1426. doi:10.1016/j.neuroimage.2010.01.075.
- Sato, M., Tremblay, P., & Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*, *111*(1), 1–7. doi:10.1016/j.bandl.2009.03.002.



- Silberstein, R. B., Danieli, F., & Nunez, P. L. (2003). Fronto-parietal evoked potential synchronization is increased during mental rotation. *NeuroReport*, *14*(1), 67.
- Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, *101*(3), 260–277. doi:10.1016/j.bandl.2007.02.008.
- Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2009). Gestures orchestrate brain networks for language understanding. *Current Biology: CB*, *19*(8), 661–667. doi:10.1016/j.cub.2009.02.051.
- Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *NeuroImage*, *25*(1), 76–89. doi:10.1016/j.neuroimage.2004.11.006.
- Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(10), 2387–2399. doi:10.1093/cercor/bhl147.
- Spencer, J. P., Perone, S., & Johnson, J. S. (n.d.). The dynamic field theory and embodied cognitive dynamics. *Toward a New Grand Theory of Development*, 86–118.
- Srinivasan, R., Bibi, F. A., & Nunez, P. L. (2006). Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency. *Brain topography*, *18*(3), 167.
- Steriade, M., McCormick, D. A., & Sejnowski, T. J. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science*, *262*(5134), 679–685.
- Towle, V. L., Bolaños, J., Suarez, D., Tan, K., Grzeszczuk, R., Levin, D. N., Cakmur, R., et al. (1993). The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy. *Electroencephalography and Clinical Neurophysiology*, *86*(1), 1–6.
- Turkeltaub, P. E., & Coslett, H. B. (2010). Localization of sublexical speech perception components. *Brain and Language*, *114*(1), 1–15. doi:10.1016/j.bandl.2010.03.008.
- Ulloa, E. R., & Pineda, J. A. (2007). Recognition of point-light biological motion: Mu rhythms and mirror neuron activity. *Behavioural Brain Research*, *183*(2), 188–194. doi:10.1016/j.bbr.2007.06.007.
- Van Der Loo, E., Congedo, M., Plazier, M., Van De Heyning, P., & De Ridder, D. (2007). Correlation between Independent Components of scalp EEG and intracranial EEG (iEEG) time series.

- van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*(3), 598–607.
- Vanni, S., Rockstroh, B., & Hari, R. (1996). Cortical sources of human short-latency somatosensory evoked fields to median and ulnar nerve stimuli. *Brain Research*, *737*(1-2), 25–33.
- Wilson, S. M. (2009). Speech perception when the motor system is compromised. *Trends in Cognitive Sciences*, *13*(8), 329–330; author reply 330–331. doi:10.1016/j.tics.2009.06.001.
- Wingeier, B. M., Nunez, P. L., & Silberstein, R. B. (2000). Spherical harmonic decomposition applied to spatial-temporal analysis of human high-density EEG. *Arxiv preprint physics/0011006*.
- Wong, P., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*, *47*(3), 693–703.
- Yang, L., Liu, Z., & He, B. (2010). EEG-fMRI reciprocal functional neuroimaging. *Clinical Neurophysiology, In Press, Corrected Proof*. doi:10.1016/j.clinph.2010.02.153.
- Yuan, H., Liu, T., Szarkowski, R., Rios, C., Ashe, J., & He, B. (2010). Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: An EEG and fMRI study of motor imagery and movements. *Neuroimage*, *49*(3), 2596–2606.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*(10), 946.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science (New York, N.Y.)*, *256*(5058), 846–849.
- Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cerebral Cortex (New York, N.Y.: 1991)*, *6*(1), 21–30.

APPENDIX. SUPPLEMENTAL FIGURES

Left  $\mu$

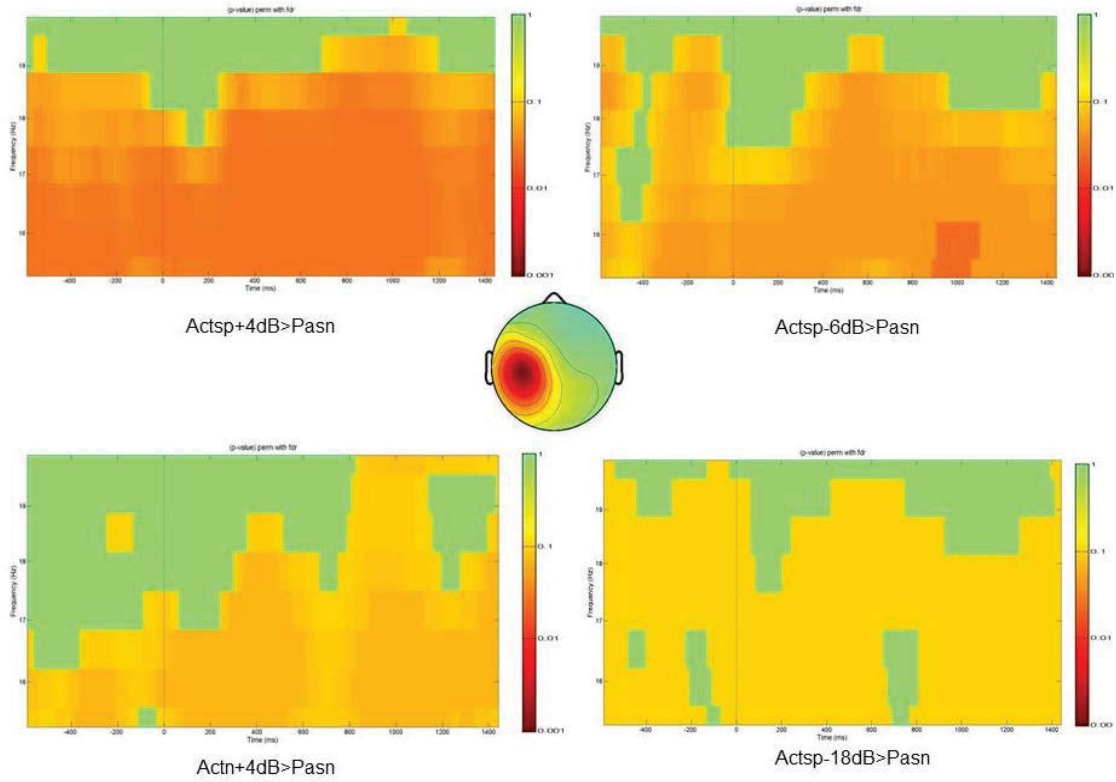
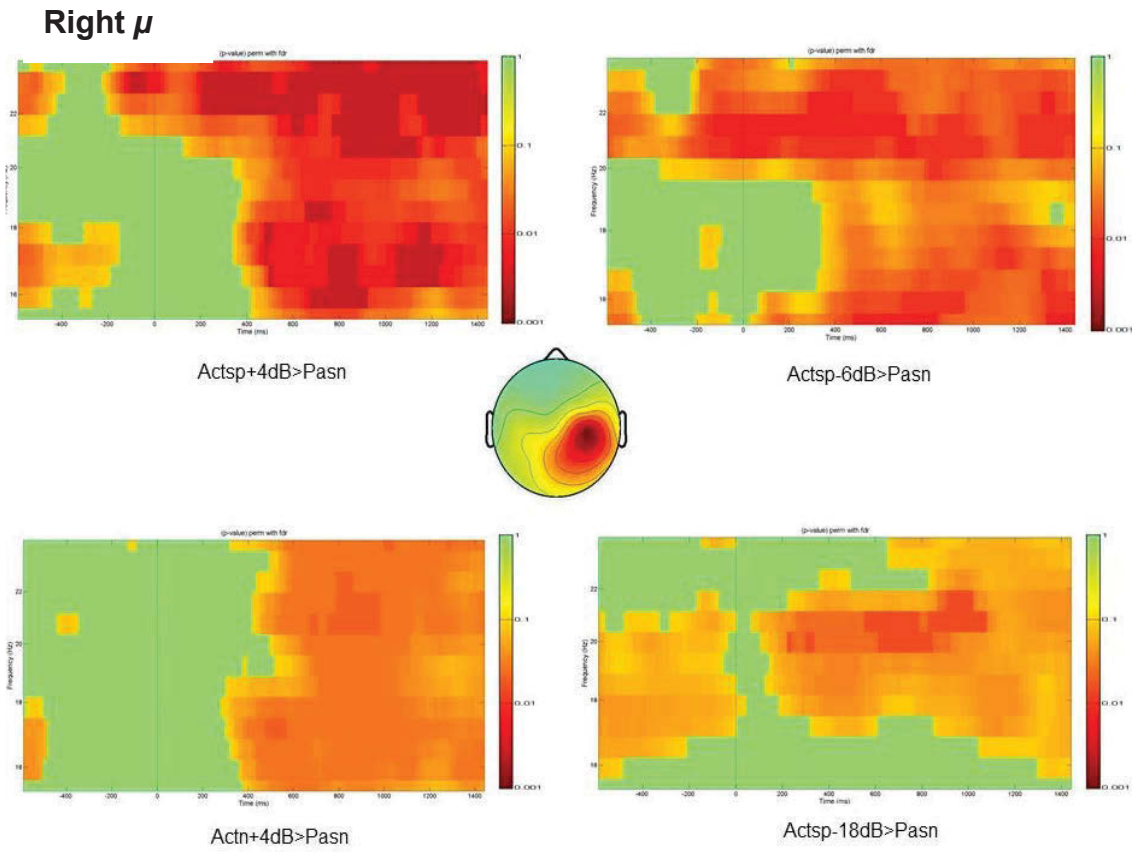
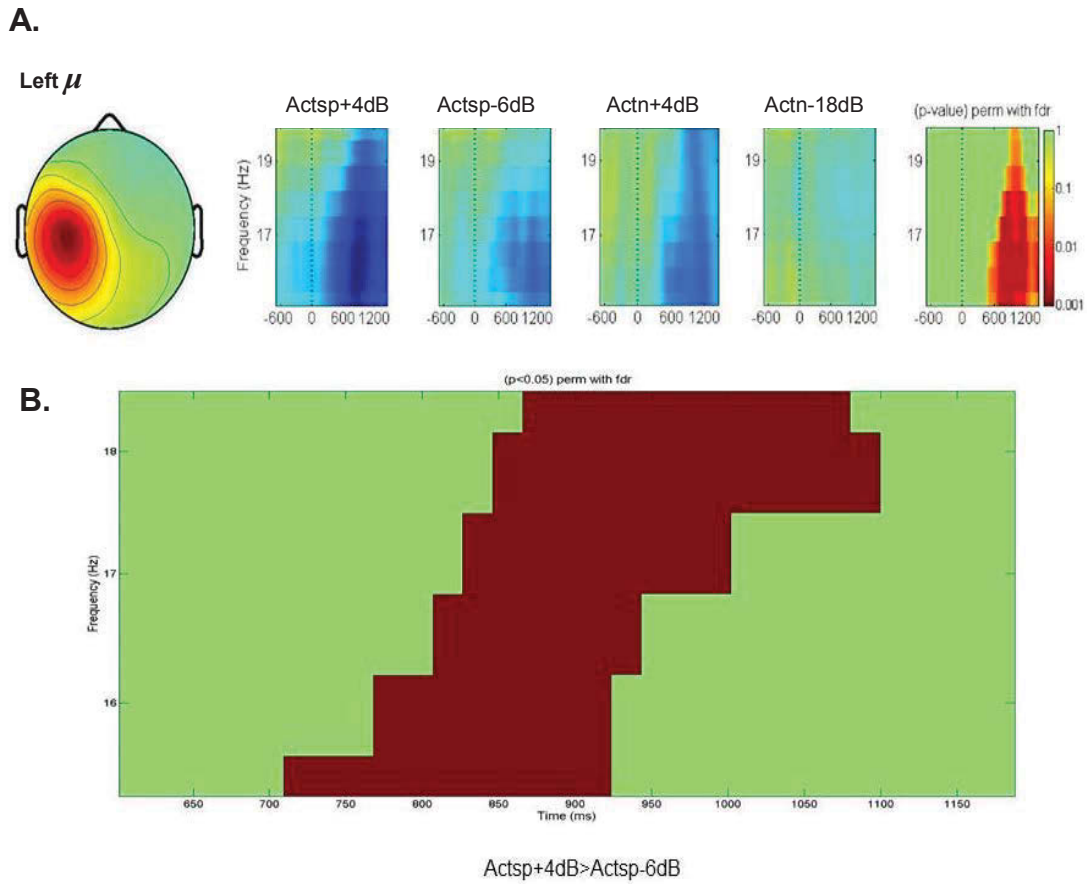


Figure A-1. Left  $\mu$   $p$ FDR in the speech and tone conditions for which an active discrimination was required.

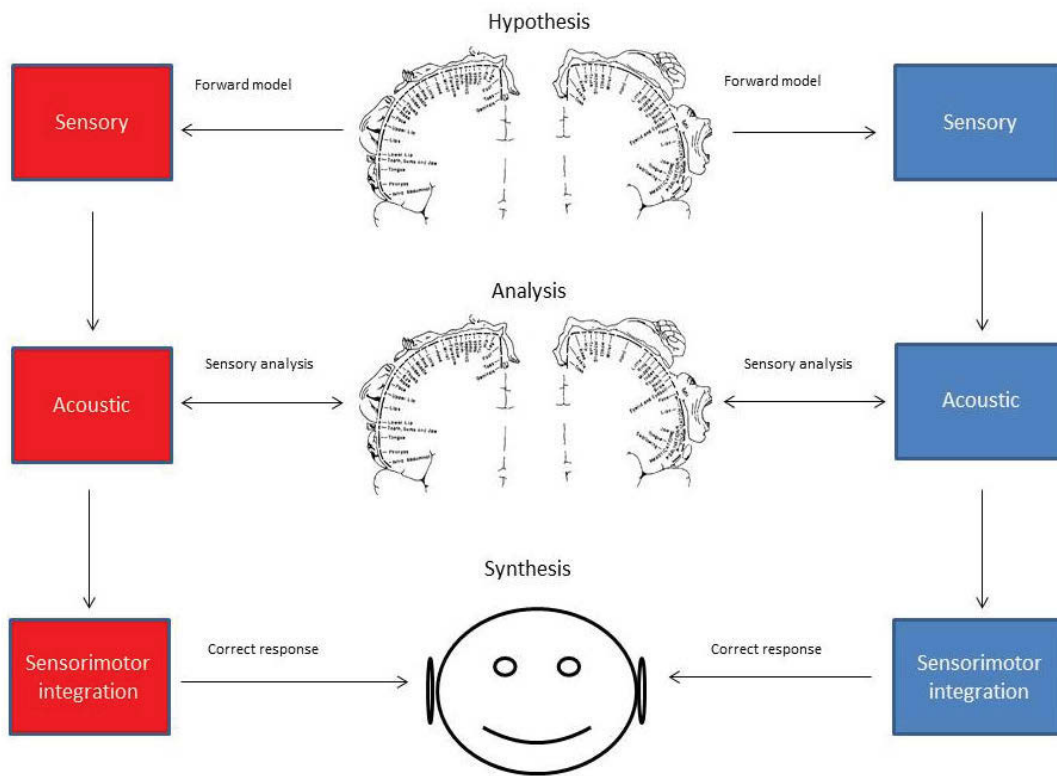


**Figure A-2. Right  $\mu$   $p$ FDR values in the speech and tone conditions for which active discrimination was required.**

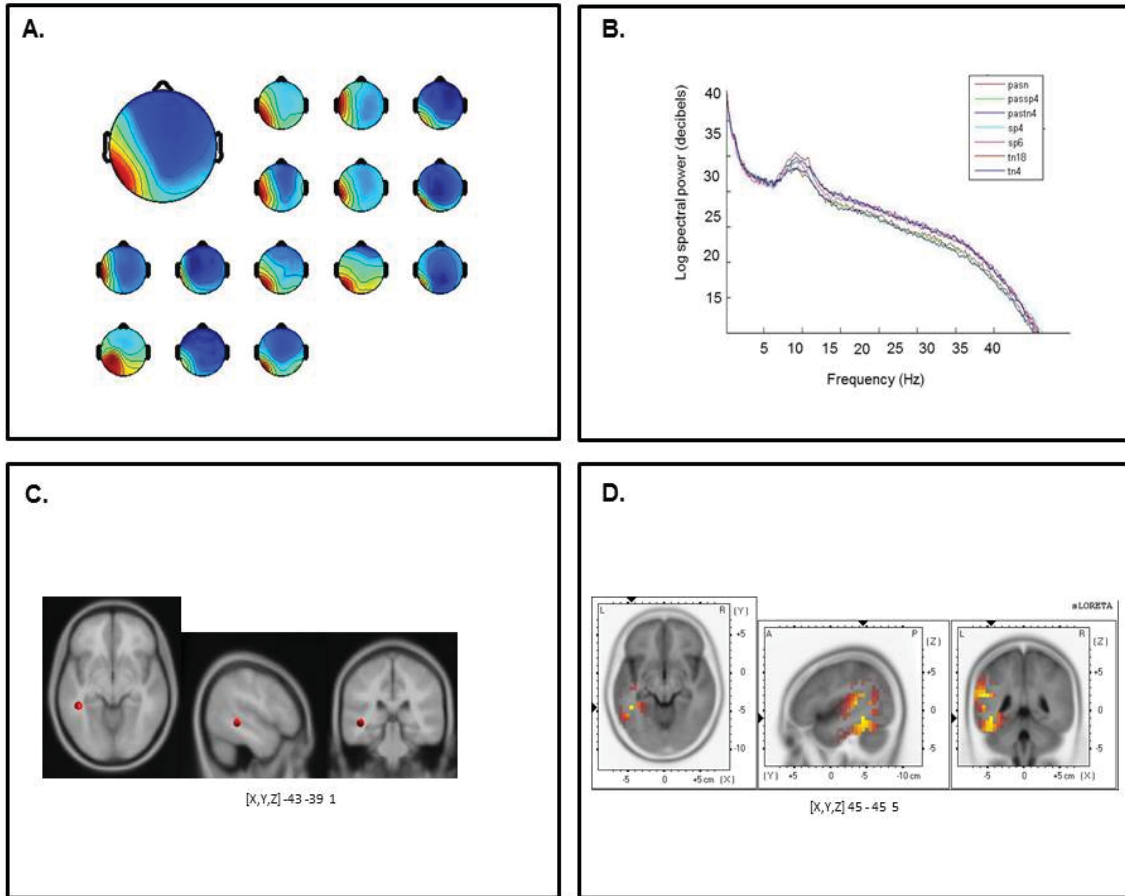


**Figure A-3. Active ERSPS for the left hemisphere  $\mu$  cluster.**

Notes. A) ERSPS in active conditions with significant differences depicted in orange and red and B) significant differences at  $p\text{FDR} < .05$  depicted in brown for the contrast Actsp+4dB > Actsp-6dB.



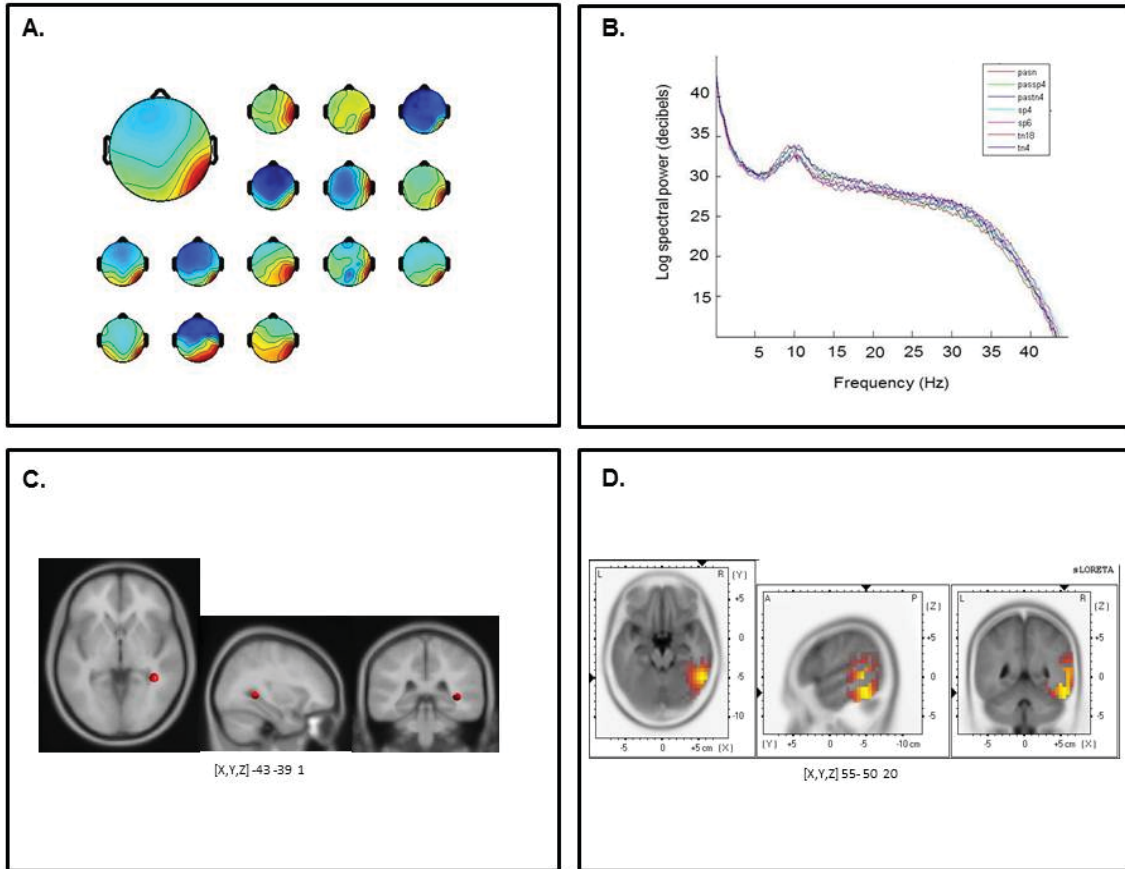
**Figure A-4. Proposed model of sensorimotor interactions in a two-forced choice discrimination task.**



**Figure A-5. Cluster results for the left-hemisphere  $\alpha$  component.**

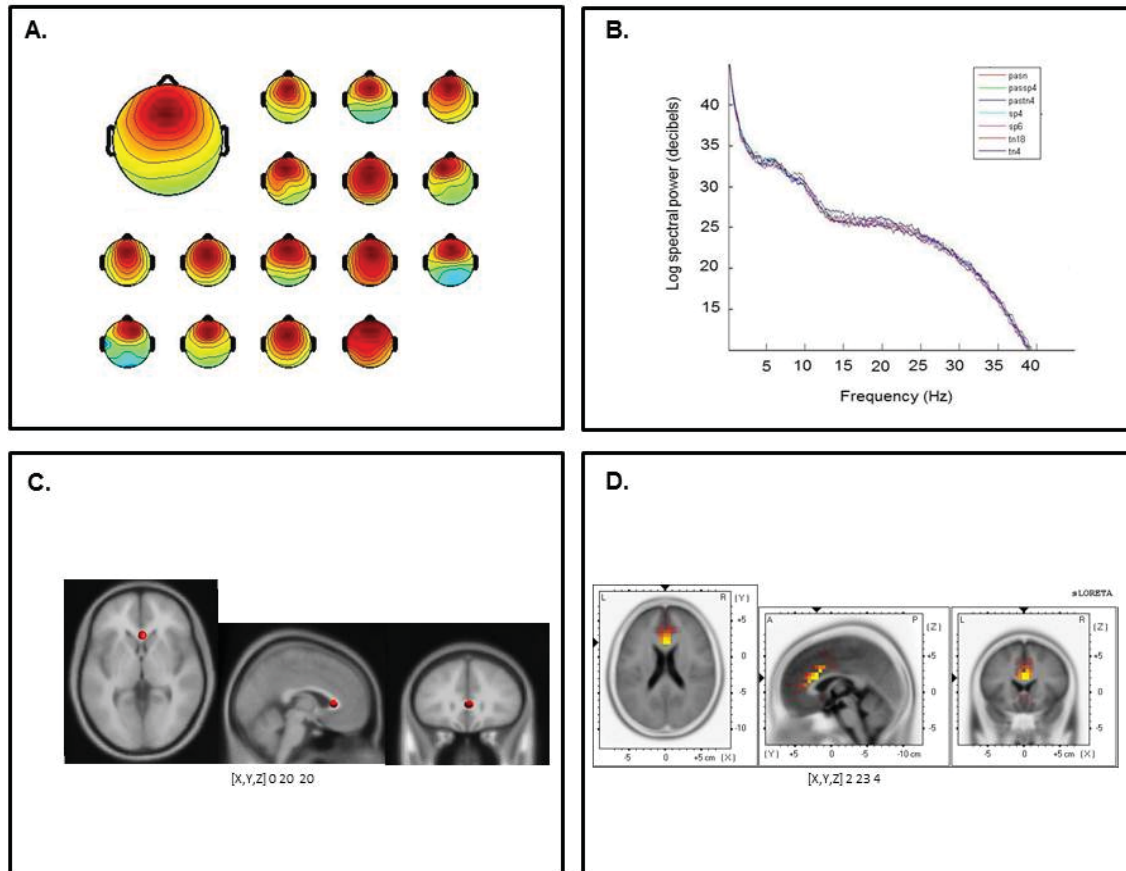
Notes. A) mean scalp potential distribution scaled to RMS microvolts ( $W^{-1}$ ), B) mean spectra of the component as a function of condition in relative decibels, C) average equivalent current dipole location (Taliarach -43,-39,5), D) maximum current source density voxels ( $p < .001$  corrected) (Taliarach-45,-25,45).





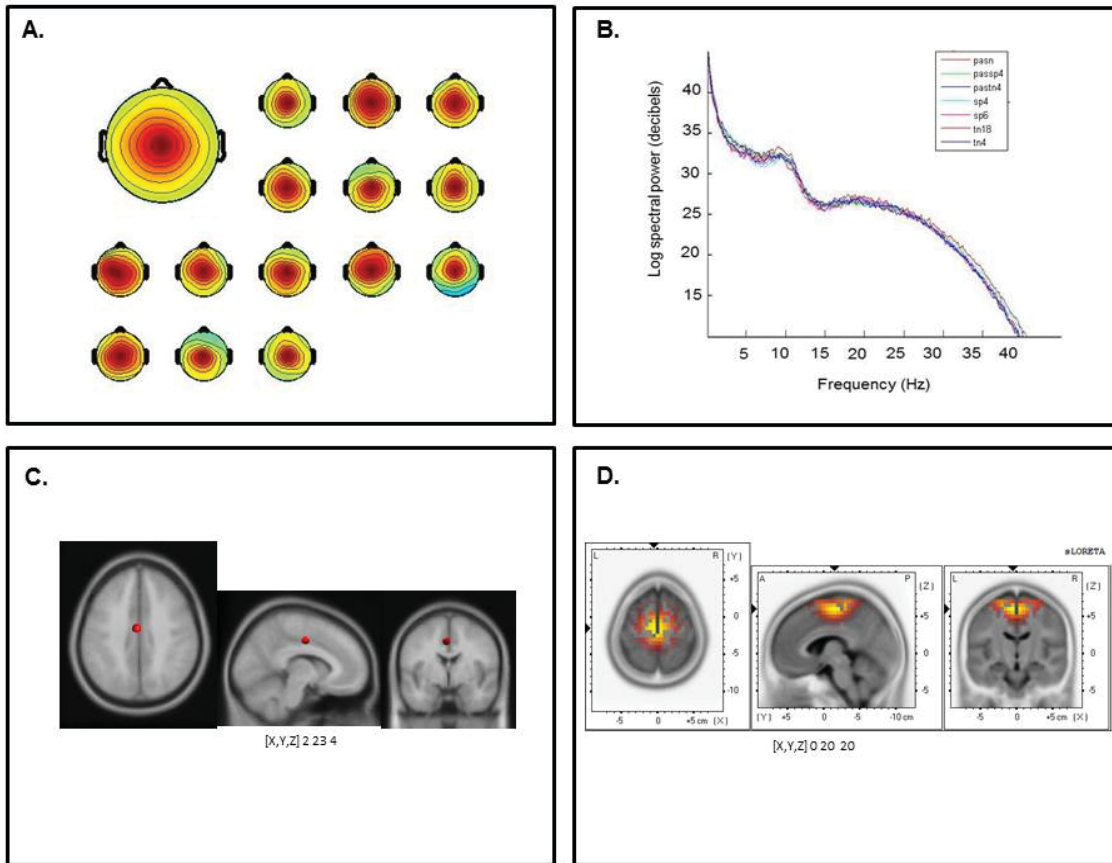
**Figure A-6. Cluster results for the right-hemisphere  $\alpha$  component.**

Notes. A) mean scalp potential distribution scaled to RMS microvolts ( $W^{-1}$ ); B) mean spectra of the component as a function of condition in relative decibels; C) average equivalent current dipole location (Taliarach 43, -39, 1); D) maximum current source density voxels ( $p < .001$  corrected) (Taliarach 55, -50, 20).



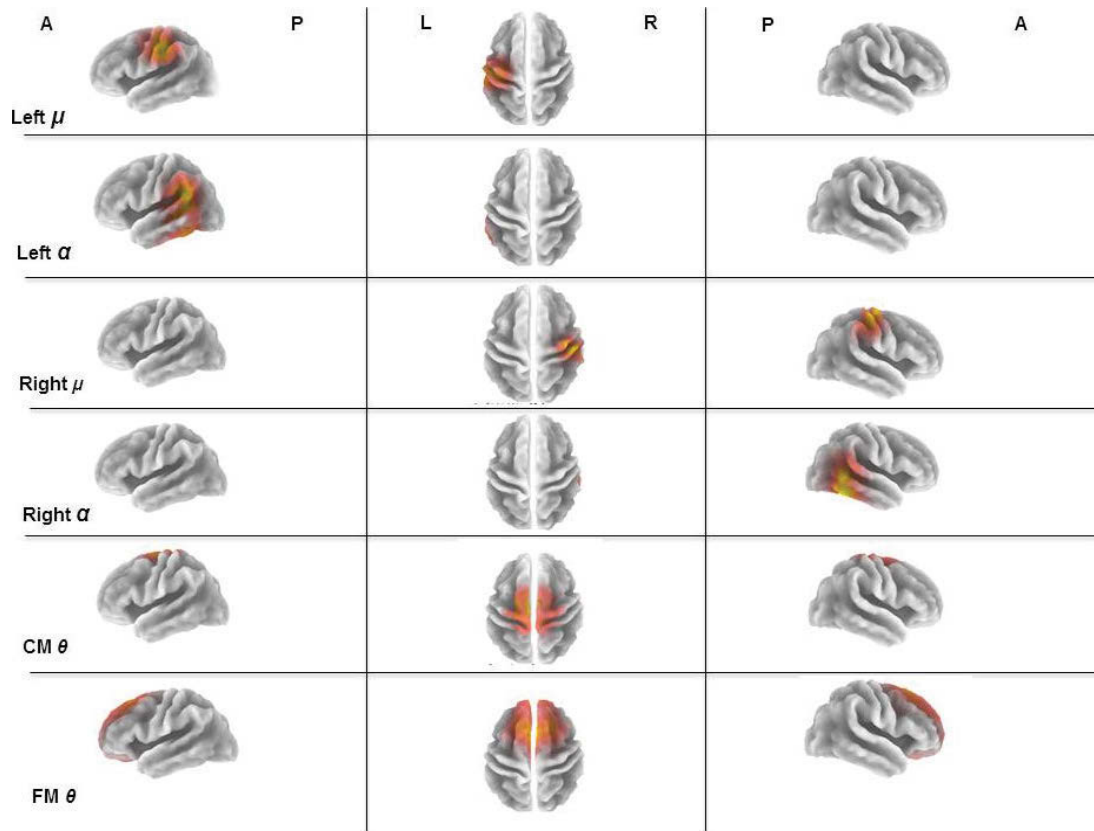
**Figure A-7. Cluster results for the frontal midline  $\theta$  component.**

Notes. A) mean scalp potential distribution scaled to RMS microvolts ( $W^{-1}$ ); B) mean spectra of the component as a function of condition in relative decibels; C) maximum current source density voxels ( $p < .001$  corrected) (Talairach 0,20,20); D) average equivalent current dipole location (Talairach 2,23,4).

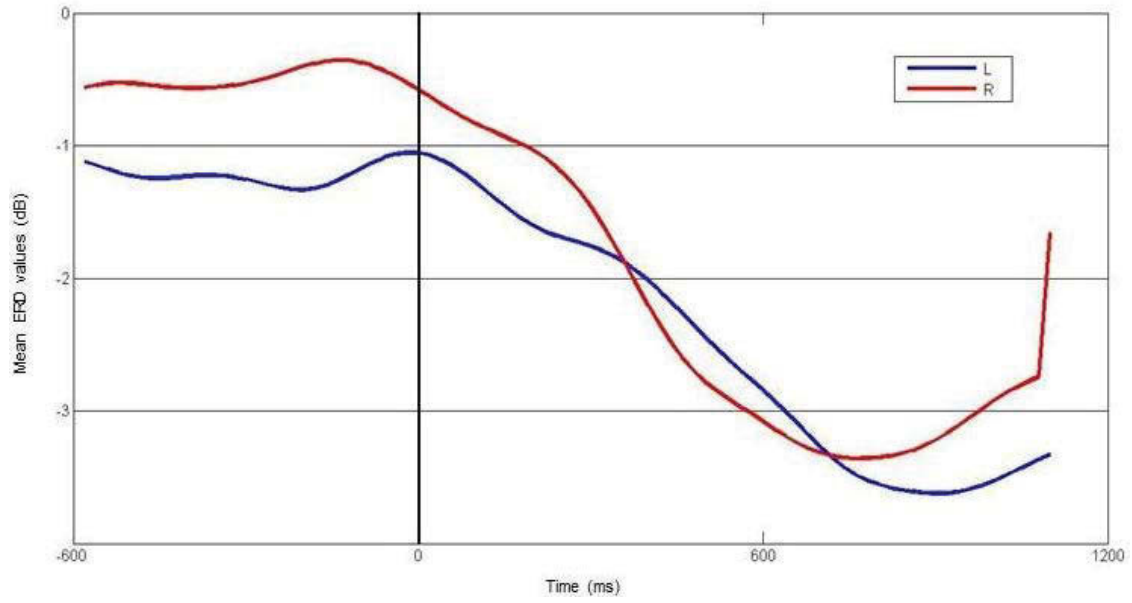


**Figure A-8. Cluster results for the central midline  $\theta$  component.**

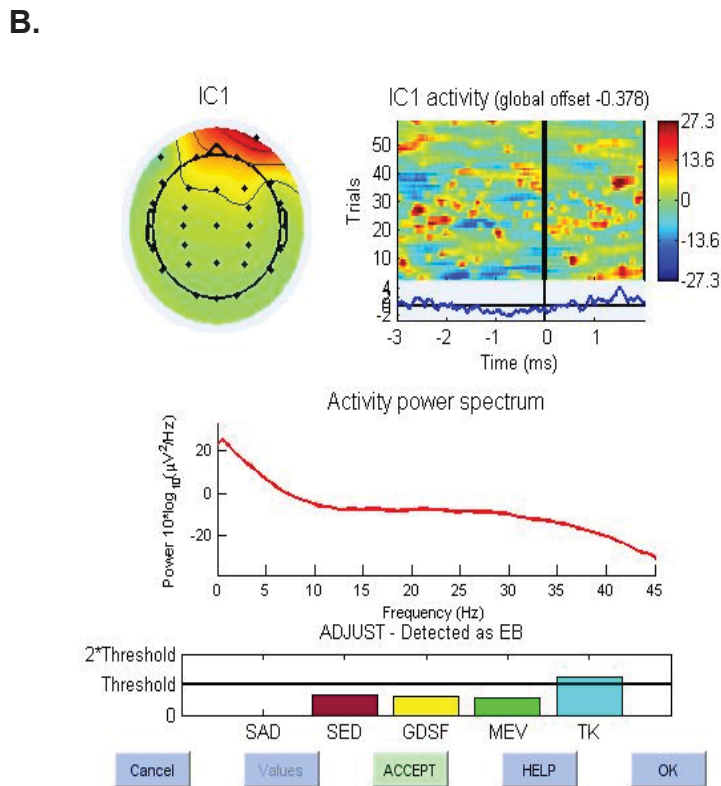
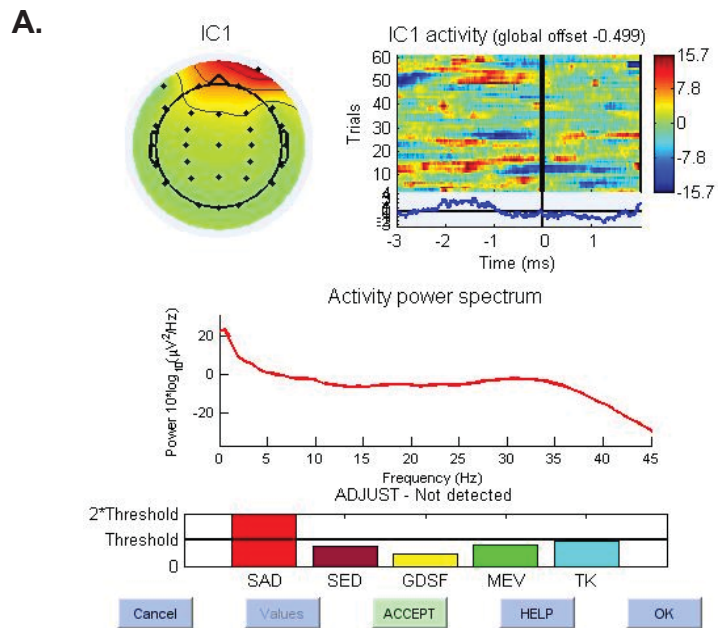
Notes. A) mean scalp potential distribution scaled to RMS microvolts ( $W^{-1}$ ); B) mean spectra of the component as a function of condition in relative decibels; C) average equivalent current dipole location (Talariach 2,23,4); D) maximum current source density voxels ( $p < .001$  corrected) (Talariach 0,20,20).



**Figure A-9. Independent component clusters labeled according to spectral signatures and depicted on a van Essen average template.**

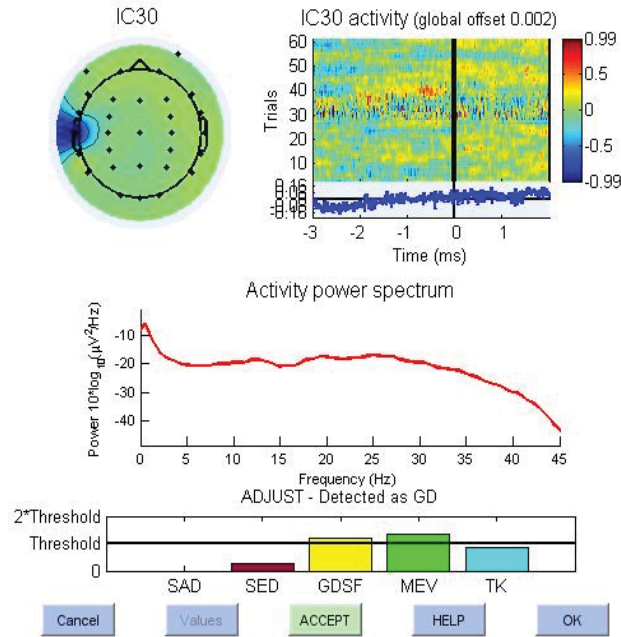


**Figure A-10.** Left and right  $\mu$  for correct trials in the Actsp+4dB condition averaged across 15-20Hz and 15-25Hz bands respectively.



**Figure A-11. Example eye-blink components with a spatial distribution near periocular channels.**

A.



B.

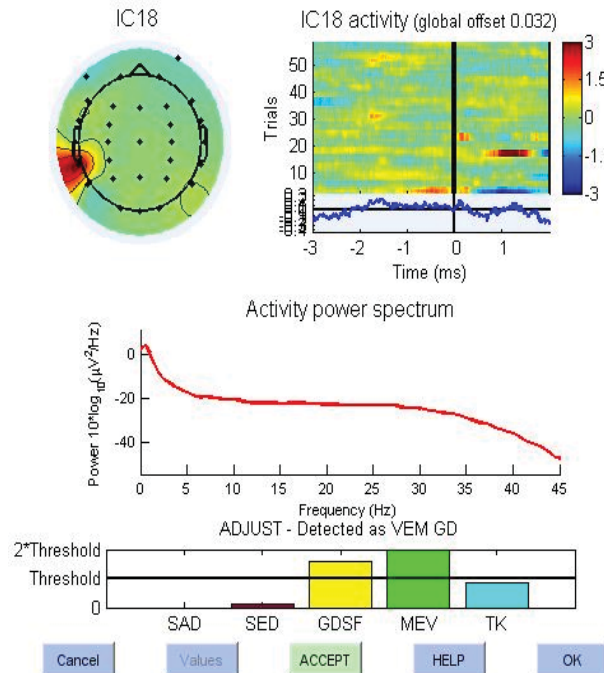
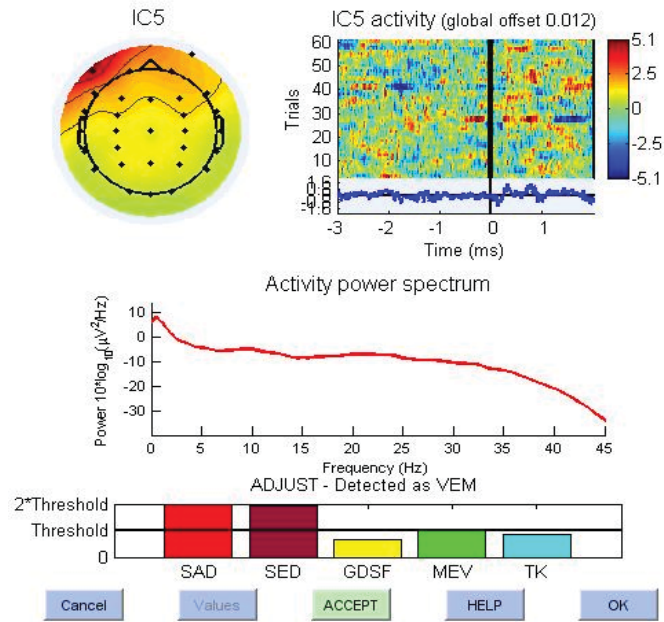


Figure A-12. Example noise components with a spatial distribution near temporal channels and spectral features consistent with temporal muscle movements.



A.



B.

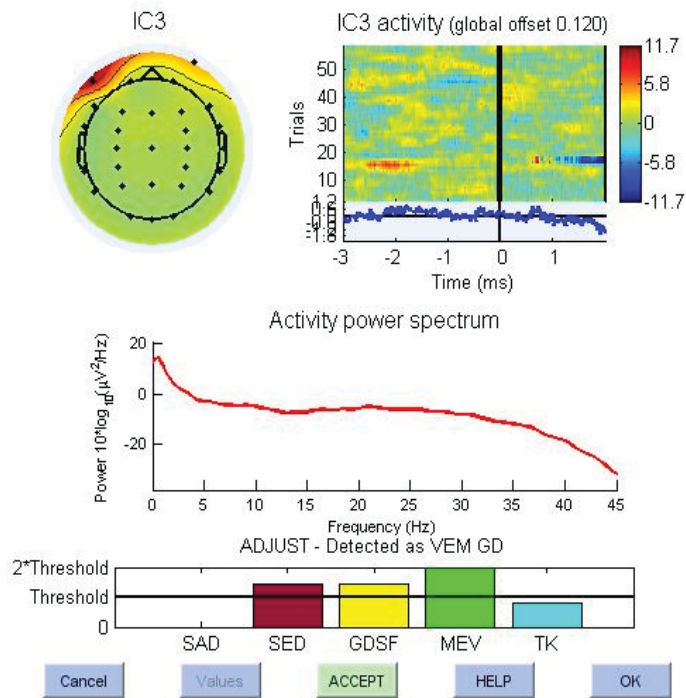


Figure A-13. Example vertical eye-movement components with a spatial distribution near periocular channel.

## VITA

Andrew Bowers IV was born in Nashville Tennessee in 1979. He received his bachelor's degree in English Creative Writing from the University of Tennessee in 2002 and his Master of Arts in Speech-Language Pathology in 2007. He received his Doctor of Philosophy from the University of Tennessee Health Science Center in 2012. Andrew has six peer-reviewed publications and 11 national conference presentations.