Abstract

AUDITORY EVENT-RELATED POTENTIALS RECORDED DURING PASSIVE LISTENING AND SPEECH PRODUCTION

by Shannon D. Swink

May 14, 2010

Director: Andrew Stuart, Ph.D.

DEPARTMENT OF COMMUNICATION SCIENCES AND DISORDERS

What is the role of audition in the process of speech production and speech perception? Specifically, how are speech production and speech perception integrated to facilitate forward flowing speech and communication? Theoretically, these processes are linked via feedforward and feedback control subsystems that simultaneously monitor on-going speech and auditory feedback. These control subsystems allow self-produced errors to be detected and internally and externally generated speech signals distinguished. Auditory event-related potentials were utilized to examine the link between speech production and perception in two experiments. In Experiment 1, auditory event-related potentials during passive listening conditions were evoked with nonspeech (i.e., tonal) and natural and synthetic speech stimuli in young normal-hearing adult male and female participants. Latency and amplitude measures of the P1-N1-P2 components of the auditory long latency response were examined. In Experiment 2, auditory

evoked N1-P2 components were examined in the same participants during self-produced speech under four feedback conditions: nonaltered, frequency altered feedback, short delay auditory feedback (i.e., 50 ms), and long delay auditory feedback (i.e., 200 ms). Gender differences for responses recorded during Experiments 1 and 2 were also examined. Significant differences were found for P1-N1-P2 latencies and for P1-N1 and N1-P2 amplitudes between the nonspeech stimulus compared to speech tokens and for natural speech compared to synthetic speech tokens in Experiment 1. These findings were attributed to differences in the spectro-temporal characteristics of the tokens. In Experiment 2, there were no significant differences in N1-P2 latencies and amplitudes across feedback conditions. To examine differences between component latency and amplitude during passive listening and active speaking, responses elicited via passively presented self-produced nonaltered and frequency altered tokens were compared to the nonaltered and frequency altered feedback active conditions. Significantly, smaller N1-P2 component amplitudes were recorded during the active versus passive speaking conditions. This finding is in accordance with research supporting feedforward and feedback theories. To further understanding of cortical processing during speech production and speech perception additional investigations are warranted in both those with normal speech and language and those with pathological speech and language, specifically, those with a speech motor disorder such as developmental stuttering.

AUDITORY EVENT-RELATED POTENTIALS RECORDED DURING PASSIVE LISTENING AND SPEECH PRODUCTION

A Dissertation

Presented to

The Faculty of the Department of Communication Disorders and Sciences

School of Allied Health Sciences

East Carolina University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy in Clinical Audiology

Presented by

Shannon D. Swink

May 14, 2010

AUDITORY EVENT-RELATED POTENTIALS RECORDED DURING PASSIVE LISTENING AND SPEECH PRODUCTION

by

Shannon D. Swink

APPROVED BY:

DIRECTOR OF DISSERTATION: _____

Andrew Stuart, Ph.D.

COMMITTEE MEMBER: _____

Sherri M. Jones, Ph.D.

COMMITTEE MEMBER: _____

Joseph Kalinowski, Ph.D.

COMMITTEE MEMBER: _____

Gregg D. Givens, Ph.D.

CHAIR OF THE DEPARTMENT OF COMMUNICATION SCIENCES AND DISORDERS:

_____

Gregg D. Givens, Ph.D.

DEAN OF THE GRADUATE SCHOOL:

_____

Paul J. Gemperline, Ph.D.

ACKNOWLEDGEMENT

First and foremost I would like to thank my dissertation director, Dr. Andrew Stuart. Words cannot express my gratitude for all the hours he has committed to helping me not only complete this work, but throughout this program. I know it was not an easy job, but he stuck with me and I am so grateful. I also want to thank the CSDI faculty and staff, especially, my committee members: Dr. Jones, Dr. Kalinowski, and Dr. Givens. Dr. Jones, thank you so much for taking the time to read my *painfully* long document and for your suggestions and ideas. Your contributions were so helpful. Dr. Kalinowski and Dr. Givens, thank you for your insightful comments and ideas. I would like to Mr. Mark Allen who helped with so many of the technical aspects of this project and allowed me to borrow so many of this tools and computers, sometimes for months. I would like to recognize family. All of you contributed in your own special way in helping me get through this program, my mother and grandmothers with their endless support and encouragement and Brody for making me laugh and giving me little special moments to hold on too, I appreciate your love more than I can say. Thank you! Last but not least I would like to thank my friends. Through the years I have met so many wonderful people that I could write another document individually thanking all of you. Given that I really do need to graduate one of these days — I cannot do this—but know I consider myself lucky to have met all of you. So thank you, Dan, Kristal, Tim, Matt, Jeff, Lauren, Nate, and Terry Hall. A special thank you to Marty—all the way to my toes — Thank you! Finally, I would like to thank my dear friend, Robyn. From day one she has been by side. We have struggled through this program/life together and we made it—for that I can be proud of both us. I would like to Thank Caleb Joyce, Graphic Designer, Austin, TX for providing the artwork presented in Figure 1 of this document.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AAF	Altered auditory feedback
ASHA	American Speech-Language and Hearing Association
ABR	Auditory brainstem response
ANOVA	Analysis of variance
BA	Brodmann Area
BF	Binaural fusion
BW	Bandwidth
CANS	Central auditory nervous system
CAP	Central auditory processing
CAPD	Central auditory processing disorder
DAF	Delayed auditory feedback
dB	Decibel
df	Degrees of freedom
DIVA	Directions Into Velocities of Articulators
DSP	Digital signal processor
EEG	Electroencephalography
η^2	Eta squared; Effect size
ERP	Event-related potential
F	*F*-ratio
F	Formant
FAF	Frequency altered feedback

FFT	Fast Fourier Transforms
FM	Frequency modulated
FS SP	Frequency shifted self-produced
HL	Hearing level
Hz	Hertz
ISI	Interstimulus interval
IWF	Individuals who are fluent
IWS	Individuals who stutter
LC	Left competing
LEA	Left ear advantage
LNC	Left noncompeting
L.Q.	Laterality Quotient
LTASS	Long term average speech spectrum
µS	Microseconds
µV	Microvolt
mm	Millimeters
ms	Milliseconds
MAF	Masked altered feedback
MEG	Magnetoencephalography
MLD	Masking level difference
MLR	Middle latency response
MMN	Mismatch negativity

MMS	Mini-Mental State
MRI	Magnetic resonance imaging
Nø	Noise in phase
Nπ	Noise out of phase
NAF	Nonaltered feedback
NF	Natural female
NM	Natural male
p	Probability
PB	Phonetically balanced
PET	Position Emission Tomography
PI	Performance intensity
pSPL	Peak sound pressure level
PTA	Pure tone average
RC	Right competing
REA	Right ear advantage
RNC	Right noncompeting
s	Second
Sø	Signal in phase
Sπ	Signal out of phase
SD	Standard Deviation
SF	Synthetic female
SL	Sensation level
SM	Synthetic male

SMA	Supplementary motor area
S/N	Signal-to-noise ratio
SP	Self-produced
SPL	Sound pressure level
SRT	Speech recognition threshold
SSI-CCM	Synthetic Sentence Identification with Contralateral Competing Message
SSI-ICM	Synthetic Sentence Identification with Ipsilateral Competing Message
SSW	Staggered Spondaic Words
STG	Superior temporal gyrus
TM	Tympanic membrane
TPP	Tympanometric peak pressure
TW	Tympanometric width
V_{ea}	Ear canal volume
VOT	Voice onset time
VOX	Voice Onset Relay
WRS	Word recognition score
Y_{tm}	Peak compensated static acoustic admittance

CHAPTER I: REVIEW OF LITERATURE

Introduction

What is the role of audition in the process of speech production and speech perception? More specifically, what is the link between the motor act of speech production and the sensory process of speech perception? In looking at the process of human communication, even in the most basic exchage, the integration between speech production and speech perception seems apparent. Historically, auditory contributions to speech production (i.e., the link between speech production and speech perception) have been discounted. Borden (1979) asserted that the auditory system is too slow for online monitoring of speech production. More recently, the advent of sophisticated electrophysiological technologies have lead to a renewed interest in examining the link between speech production and speech perception (Guenther, Ghosh, & Tourville, 2006) and to the constructing of more contemporary models of speech production, which account for auditory system involvement.

One such model posits that speech production and speech perception are intimately linked via feedforward and feedback control subsystems working in concert. These feedforward and feedback control system facilitate communication and importantly allow for the monitoring of internal and external vocalizations (e.g., Frith, 1992; Guenther, 2006, 2007; Jeannerod, 1998; Numminen, Salmelin, & Hari, 1999). In brief**,** it has been postulated that at the level of the feedforward control system, the motor program for speech is generated (i.e., the intention to speak is formulated and the neural discharge for such is programmed). Simultaneously, an efference or sensory copy (Jeannerod, 1998) of the motor program is also generated (Sherrill et al., 2002). This efference copy produces a corollary discharge signal (Sperry, 1950) based on internal models (Kawato & Wolpert, 1998) of previous speech acts. The corollary discharge

provides a representation of the expected sensory input (i.e., expected auditory or somatosensory feedback) from the vocalization. Within, the feedback control system, a comparison is made between the spoken message (i.e., actual feedback) and the corollary discharge (i.e., expected feedback; von Holst, 1954). If the two signals match (i.e., there is not a discrepancy), neuronal suppression will ensue as a result of the net cancellation of the "excitatory input from the auditory periphery" (Guenther, 2007, p.36), and forward flowing speech will occur. In other words, corollary discharge will cancel or inhibit the incoming excitatory auditory input, which results in an overall reduction of neural activity within the temporal lobe (Ford, Mathalon, Heinks, Kalba, & Roth, 2001; Katahira, Abla, Masuda, & Okanoya, 2008, von Holst, 1954).

If a mismatch occurs between the expected auditory feedback and the actual auditory feedback during the act of speaking (i.e., a speech error is produced), the corollary discharge will not match the actual feedback, the incoming auditory information will not be cancelled (i.e., neural activity in the temporal lobe will not be suppressed), which result in the disruption of free flowing speech (Ford, Mathalon, Kalba, Whitfield, Faustman, & Roth, 2001; Ford & Mathalon, 2004; Guenther, 2001, 2006, 2007; Guenther & Ghosh, 2003; Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Heinks-Maldonado, Nagarajan, & Houde, 2006; Heinks-Maldonado, Mathalon, Houde, Gray, Faustman, & Ford, 2007). Overtly these disruptions may appear as abrupt interruptions, where the speaker stops speaking immediately (i.e., mid-sentence or word) to correct the error (Levelt, 1983, 1989; Nooteboom, 1980) or a more subtle interruption, where the speaker continues speaking until the end of their statement and then subsequently interjects the correction (Seyfeddinipur, Kita, & Indefrey, 2008). The latter more subtle type of interruption provides the speaker with a time window in which they can formulate an appropriate

correction (Blackmer & Mitton, 1991) without severe consequences to the flow of speech (Hartsuiker, Pickering, & de Jong, 2005).

According to the concepts of feedforward and feedback control systems, auditory monitoring allows a speaker to recognize disruptions or speech errors so that corrective self-repairs can be initiated. At the cortical level, once a speech error is detected, the unsuppressed excitatory input initiates activation of corrective motor commands (Guenther, 2007). In additions, auditory monitoring differs depending on the locus of incoming auditory information (i.e., internally generated signals are monitored differently than externally generated signals). Upon hearing an externally generated vocalization, cells within the auditory cortex are activated due to the absence of a corollary discharge signal (Aliu, Houde, & Nagarajan, 2008), in turn, the process of sensory to motor matching does not occur, and the signal is deemed as externally generated. These signals are then monitored for features such as intent and context. Interestingly, the monitoring of external auditory input occurs concurrently with the process of auditory self-monitoring in such a way that speech production does not abolish perception.

The process of monitoring via corollary discharge matching is not exclusive to the human species. Numerous researchers have posited the presence of corollary discharge systems across various animal models including crickets (Poulet & Hedwig, 2003a, 2003b, 2006), echolocation bats (Metzner, 1989; 1993; Schuller, 1979; Suga & Schlegal, 1972), non-human primates (Müller-Preuss, 1978; Müller-Preuss, Newman, & Jürgens, 1980; Müller-Preuss & Ploog, 1981), and song birds such as the zebra finch (Cooper & Goller, 2004; Iyengar & Bottjer, 2002; Leonardo & Konishi, 1999; Mooney & Prather, 2005; Shea & Margoliash, 2003). The ability to distinguish between internal versus external auditory inputs, especially during vocalizing, would seem to provide an evolutionary benefit against dangerous situations and attacks from predators.

Electrophysiological measures and brain imaging have provided noninvasive means to examine feedforward and feedback theories and the effect of sensory to motor matching. In the following review, summaries of the literature pertaining to speech production and perception with respect to auditory system involvement, electrophysiological auditory assessment with a focus on gender differences, brain imaging, and indices of sensory-to-motor integration will be presented. A model of auditory self-monitoring will also be proposed and the relevant research questions introduced.

An Overview of Theories for Speech Production and Speech Perception

The motor act of speaking is a highly complex action that requires the integration and interconnection among the frontal, temporal, and parietal lobes of the brain (Greenlee et al., 2004). Guenther (2006) posits that "there are three main types of information involved in the production of speech sounds: auditory, somatosensory, and motor, which are represented in the temporal, parietal, and frontal lobes of the cerebral cortex, respectively" (p. 351). The act of speaking first requires the conceptual formulation of a message (Levelt, 1989) followed by the encoding of this message into a sound particular sequence that is understood by communicative partners. Although, the act of speech production is a complex, most individuals are able to encode the intention or desire to speak and subsequently perform the overt act of speaking in a rapid and precise succession (i.e., within in hundreds of milliseconds (ms); Munhall, 2001).

As presented above, the question can be proposed: What role does audition play in the process of speech production? With respect to communication, the overall goal of speech is to relay a sequence of sounds and linguistic information that can be encoded into an understandable message. Numerous languages are spoken across the globe. Each language having its own grammatical structure and set of rules governing how these vocalized sounds are sequenced.

During infancy and childhood, one's particular linguistic sequence (i.e., native language) is formulated through the act of hearing their native dialect (Kuhl, 2000). Given this, again, it is logical to assume there is a strong relationship between speech production and speech perception (Peschke, Ziegler, Kappes, & Baumgaertner, 2009).

Effective communication also depends on one's ability to monitor self and externally produced vocalizations. Though, the exact relationship between speech production and perception is unclear, auditory feedback has been shown to play a crucial role in the monitoring of both internally and externally generated vocalizations (Tourville, Reilly, & Guenther, 2008). For example, speakers adjust the amplitude and/or pitch of their voice when there is a change in the surrounding noise levels or when perceived auditory feedback is altered (Lane & Tranel, 1971; Howell & Sackin, 2002; Purcell & Munhall, 2006; Xu, Larson, Bauer, & Hain, 2004).

It was posited that speech production and perception are mediated by feedforward and feedback control systems. As defined by Tourville et al. (2008), feedback control occurs when "task performance is monitored during execution and deviations for the desired performance are corrected according to sensory information" (p. 1429) and feedforward control occurs when "task performance is executed from previously learned commands, without reliance on incoming task-related sensory information" (p. 1429). Accordingly these authors suggest that speech production relies on the integration of feedforward and feedback control subsystems, both of which are affected by the perceived auditory feedback. Some, however, discount any involvement of the auditory system. Following is a review of several theories and models both supporting and refuting auditory system involvement in the production of speech, including the "Directions Into Velocities of the Articulators (DIVA)" model, corollary discharge theories, the "Motor Speech Theory" and the "Perceptual Loop" theory.

Theories Discounting Auditory System Contributions During Speech Production

Historically, the auditory system has been discounted as contributing to the process of speech production. Borden (1979) concluded that speech monitoring is not controlled by the auditory system's utilization of auditory feedback, for two reasons. First, she reasoned that auditory feedback is not essential for the production of intelligible speech because those who are postlingually deafened generally do not experience breakdowns in speech production. Second, Borden concluded that the auditory system does not process rapidly occurring speech signals fast enough for the online correction of speech. Therefore, the correction of a speech error would be initiated after the utterance is spoken.

Levelt (1983; 1989) also discounted the auditory system, specifically auditory feedback, as crucial for speech production. His "Perceptual Loop Theory" proposed that speech production is monitored via an "inner monitoring loop" that monitors the phonetic plan (i.e., the intended spoken message or inner speech) prior to vocalization, thus, allowing for the detection of errors before the error is spoken. Once an error is detected speech production is halted or interrupted and the error corrected. Levelt theorized the speech comprehension system regulates the monitoring of self-produced speech in the same manner as external speech. In that, the same mechanism that allows for the understanding of incoming speech also examines the phonetic plan for errors. In terms of the temporal aspects of this model, Levelt (1989) proposed that the phonetic plan is analyzed and errors detected within 150 to 200 ms post formulation. Articulation does not precede until 200 to 250 ms, therefore, the 100 to 150 ms time window between detection and articulation would allow the error to be corrected before it was spoken. In light of findings reported by Wheeldon and Levelt (1995), Levelt, Roelofs, and Meyer (1999) revised the perceptual loop theory, expanding the function of the inner loop to encompass not only phonetic

plan monitoring, but also suggesting that the inner loop is responsible for assessing preceding phonemics as well.

Howell and Sackin (2002) concluded that auditory feedback is not necessary for speech monitoring because the spoken output is presented to the speaker via bone conduction. For auditory monitoring to take place the speaker must receive a vertical copy of the speech utterance. This vertical copy provides the speaker with information regarding vocalizations so the occurrence of speech errors can be noted and addressed. However, when speech is produced articulatory structures generate internal noise that is transmitted via bone conduction to the cochlea. The bone-conducted noise arrives simultaneously and at the same intensity as the air conducted auditory signal. This simultaneous presentation causes the air-conducted signal, containing linguistic cues for message decoding, to be masked by the bone-conducted signal, which mostly contains the fundamental frequency of the speaker's voice and structural (e.g., the skull) resonances (Howell & Powell, 1984). This masking effect degrades the amount of useful linguistic information provided by the air-conducted signal conflicting with the notion that speech is monitored via auditory feedback.

Theories Supporting Auditory System Involvement During Speech Production

Numerous investigations have been performed looking at speech production and speech perception independently; however, relatively few models/ theories proposed consider these two processes collectively (Peschke et al., 2009). One of the most well known theories considering a perception and production link is the "Motor Theory of Speech Perception" (Liberman, 1957, Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985). The Motor Theory of Speech Perception supports the following core assumptions: speech perception is the result of processing of phonetic or articulatory gestures, speech processing is special in

terms of how speech is processed in the auditory system, and speech perception and production are intimately linked as speech production maps auditory perception (Liberman et al., 1967; Liberman & Mattingly, 1985; 1989; Mattingly & Liberman, 1988).

Feedforward and feedback control theory. Another model is the DIVA model of speech production, devised by Guenther and colleagues. DIVA is a functional neuroanatomical and computational model that incorporates both feedforward and feedback control subsystems in the explanation of speech production and speech perception (Guenther, 1994, 1995, 2007). As suggested by the title, "a central aspect of the model concerns how the brain transforms desired movement directions in sensory space into velocities of the articulators…"(Guenther, 2007, pp. 33-34). The DIVA model also highlights the interconnections between cortical motor, auditory, and somatosensory areas (Guenther, 1994; 1995; 2007; Guenther et al., 2006; Guenther, Hampson, & Johnson, 1998; Tourville et al., 2008). Interestingly, theorized concepts of this model have been applied and investigated mathematically and computationally using a computer interface and articulatory synthesizer (Maeda, 1990). An articulatory synthesizer is a model of the human vocal tract and cortical areas that uses information based on the "articulator positions" (Guenther, 2007, p. 8) to produce sounds. The process of speech production is computationally defined via the use of mathematical algorithms and computer manipulations to simulate speech production at the level of the articulators (Stevens, 1998) and/or at the cortical and subcortical levels (Guenther, 2006). For example, computationally the DIVA model has been investigated utilizing a computer interface and synthesizer, which has the ability to learn and produce speech sounds such as words, syllables, and phonemes (Guenther, 2007). These computer simulations are then compared to anatomical locations "specified in the Montreal Neurological Institute coordinate frame" (Guenther & Vladusich, 2009, p. 5) in order to associate the computational

data with specific anatomical structures. Refer to Guenther (2006, 2007), Guenther et al. (2006) or Guenther and Vladusich (2009) for a schematic outline of the DIVA model.

According to the DIVA model, speech is initially formulated in the "speech sound map cell" thought to lie within the left ventral premotor areas and the posterior Broca's area in the left frontal operculum (Guenther, 2007; Tourville et al., 2008). Guenther (2007) defines a "map" as the grouping of neurons involved with particular stages of speech production. He defines a "speech sound" as a "phoneme, syllable, or word" (Guenther, 2007, p. 34). The "speech sound map" therefore is the group of neurons involved in the initiation and formulation of speech sounds. The "speech sound map cell" is the particular cell within the map that represents the individual speech sounds - "each speech sound is represented by a different cell in the model's speech sound map" (Guenther, 2007). Once the speech sound map is initiated, feedforward control subsystems are utilized to send "motor commands" or "feedforward commands" to "articulatory maps" in the primary motor cortex. The feedforward commands subsequently regulate and initiate action before the actual motor act is preformed. Anatomically, Guenther (2007) suggests that feedforward commands are controlled by the synaptic "projections from the left frontal operculum to the primary motor cortex, supplemented by cerebellar projections" (p. 36). At the level of the articulatory maps, located within the primary motor cortex, the motor plan regulating the motor movements and sequencing needed to produce speech sounds is encoded and synaptic projections are sent to the articulators—the lips, tongue, and vocal folds. Thus, feedforward commands initiate the overt production of speech sounds, which in turn are processed as sensory feedback activating the feedback control subsystem.

Guenther (2007) proposed that both auditory feedback and somatosensory feedback are encompassed under the feedback control subsystem. Both types of feedback are essential in the

learning and production of speech sounds. With regards to the auditory feedback control subsystem, "axonal projections from the speech sound map cells in the left frontal operculum areas to the higher-order auditory cortical areas embody the auditory target region for the speech sound currently being produced" (Guenther, 2007, p.35). Auditory target regions are the sensory representations of the expected auditory feedback produced from self-vocalizations (i.e., what the individual expects to hear when they speak). According to the DIVA model, the target region accounts for some variability within the auditory signal; however, if the actual auditory feedback varies too drastically from the auditory target region the "auditory error map" is activated and measures to self-correct the error ensue (Guenther et al., 1995). It is hypothesized that auditory error map cells are located in the posterior superior temporal gyrus and the planum temporale. At the level of the auditory error map the *expected* auditory feedback (i.e., the auditory targets) and the *actual* auditory feedback are compared. If the actual feedback and the expected feedback do not match, increased activation within the ventral motor, premotor, and superior cerebellar cortex (Tourville et al., 2008), signals that a speech error has occurred. Projections from the auditory error map along with projections from cerebellum drive the production of motor commands (Ghosh, 2004) used to correct the speech error through

If the actual auditory feedback and the auditory target region do match, then there is an inhibitory effect on neurons within the auditory error map. In that, the inhibitory characteristics from the projections within the speech sound map to the auditory cortex cancel excitatory inputs from the neurons within the auditory error map. Somatosensory feedback control is postulated to function in the same way as auditory feedback control (Guenther, 2006, 2007). A "somatosensory state map" is generated within the postcentral gyrus and the supramarginal gyrus (Guenther, 2007). Tactile and proprioceptive feedback produced at the level of the articulators is

compared to the "somatosensory error map" located within the supramarginal gyrus. If somatosensory feedback is not within the somatosensory target region, the motor cortex is activated via synapses within somatosensory error map and corrective commands are produced (Guenther, 2007).

The DIVA model has also been utilized to explain speech and language development. During the initial stages of language development, speech production is highly reliant on the feedback control subsystems and feedback commands for the tuning of these feedforward commands (Guenther, 2007). Hypothetically, new speech sound maps and auditory targets are generated from hearing spoken language. During the babbling stage the child attempts to produce vocalizations that match the previously constructed speech sound maps. Initially, the infant is unable to reach the auditory target and proceeded in the production of speech errors. When an error is produced, the auditory or somatosensory error maps are activated, and corrective motor commands generated. These corrective commands tune the feedforward commands (i.e., the motor plan regulating the articulators) and are stored for future productions within the feedforward control system.

Once the desired target regions are reached and the feedforward commands are tuned, the speech production process transitions from a feedback controlled act to a feedforward driven act. In other words, as the feedforward commands are tuned via feedback commands they become more adapt at producing speech sounds reducing the number of produced errors. As the number of produced errors is reduced, one's reliance on the feedback control system is also reduced. At this point the feedforward control subsystem is sufficiently capable of regulating speech production (Guenther et al., 2006).

As an aside, Guenther and Vladusich (2009) reasoned that the speech sound map, regulating feedforward control, is comprised of neurons displaying "mirror neuron" characteristics. Mirror neurons are dual function neurons that allow firing when performing an action and when observing and action performed by another (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003; Gallese, Fadiga, Fogassi & Rizzolatti, 1996; Kohler, Keysers, Umilta, Fogassi, Gallese, & Rizzolatti, 2002). Originally, mirror neurons posited to be found within the primate premotor cortex (di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992). The primate premotor cortex is suggested to be equivalent to the Brodmann area 44 of Broca's area in humans (Rizzolatti & Arbib, 1998). Guenther and colleagues asserted that mirror neurons located within Broca's area are crucial to the learning of speech via an imitation process (Iacoboni & Dapretto, 2006). In other words, language development is related to the encoding, via mirror neurons, of the motoric actions needed to produce speech.

Corollary discharge theory in humans. Although conceptually similar to ideas postulated by Guenther and colleagues regarding the matching between the actual feedback and the excepted feedback, the notion of "efference monitoring" (Postma, 2000) or "corollary discharge" also provides a link between speech production and perception. The presence of corollary discharges has been noted across multiple motor-to-sensory domains including the visual domain (Welch, 1978) and somatosensory domain (Blackmore, Rees, & Frith, 1998).

According to the corollary discharge theory, motor and sensory events are related through the production of corollary discharges and the comparison of these corollary discharges with actual sensory events. Corollary discharge is the prediction of self-produced sensory input as a result of the motor command (Feinberg, 1978). In other words, an individual formulates motor commands for motor movements (e.g., speech) in the frontal lobe. During speech production, a

motor command is projected from the frontal lobe to the temporal lobe; an "efferent copy" is created from internal or stored models of previous motor acts (e.g., speech motor acts). The efferent copy then produces a corollary discharge representation of the expected sensory input.

When the speech command is executed, auditory feedback of the spoken message is heard at the level of the peripheral auditory system and processed through the central auditory pathway to the auditory temporal lobe. A comparison is made between the spoken message (i.e., auditory feedback) and the corollary discharge (i.e., predicted input). If a match occurs (i.e., there is no discrepancy) between the two signals, there will be a net cancellation of the cortical potentials, ultimately resulting in a dampening of cortical activity (Von Holst, 1954). This cancellation produces forward flowing speech through the suppression of auditory cortical activity within the auditory cortex and the auditory signal is perceived as self-produced.

If the signals do not match (i.e., there is a discrepancy) or if corollary discharge is absent, a net cancellation of cortical activity does not occur, the sensory event is magnified, and the auditory signal is perceived as externally generated (Ford, Mathalon, Heinks et al., 2001; Ford & Mathalon, 2004; Heinks-Maldonado et al., 2005; Heinks-Maldonado et al., 2006; Heinks-Maldonado et al., 2007). Hence, during speech production the function of corollary discharge is twofold: motor to sensory matching allows for the monitoring of self-produced vocalizations and allows self-generated vocalizations to be distinguished from externally generated vocalizations (Ford et al., 2002).

Corollary discharge theory in animal models. The presence of corollary discharge systems has been demonstrated across various animal species. It has been suggested that corollary discharge mechanisms provide a method for monitoring self-generated motor functions and allows for the discrimination between self-generated actions and externally generated

sensory stimuli (Frith, 1992; Sperry, 1950). Crapse and Sommer (2008) suggested a corollary discharge mechanism is present in all species of animals; however, the type and function of corollary discharges varies in accordance with the specific needs of the animal. They classified corollary discharge as either "lower-order corollary discharge" or "higher-order corollary discharge." Theoretically, lower-order corollary discharges are responsible for regulating lower motor to sensory commands, whereas, higher-order corollary discharges are responsible for regulating higher and more complex motor to sensory commands. For example, lower-order corollary discharge "signaling is used for the functions such as reflex inhibition and sensory filtration" (Crapse & Sommer, 2008, p.588) and higher-order corollary discharge "signaling participates in functions such as sensory analysis and stability, and sensorimotor planning and learning" (Crapse & Sommer, 2008, p.589).

In those animals capable of producing sound, it is essential that the ability to detect incoming externally generated sounds, especially those indicate approaching danger, is not impeded by internally generated sounds. One such animal model is the stridulating cricket. Crickets rub their forewings together (Poulet & Hedwig, 2006) to produce sounds or chirps at levels as high as 100 dB SPL (Jones & Dambach, 1973). This process is referred to as stridulation (i.e., singing). Located approximately 50 mm from the anatomical point of stridulation (i.e., the forewings), on the forelegs, lies the cricket's tympanic membranes, which are used for the detection of incoming sounds (Hoy & Robert, 1996). Given, the anatomical proximity of the cricket's tympanic membranes and forewings it was initially reasoned that self produced stridulations should mask the detection of other auditory signals. Heiligenburg (1969) and Jones and Dambach (1973) provided behavioral evidence to the contrary in showing that crickets are able to process external auditory input during overt stridulations.

Poulet and Hedwig (2002, 2003a, 2003b, 2006) and Poulet (2005) proposed that a corollary discharge mechanism acts to inhibit sensitivity to self-produced stridulations at the level of the central nervous system while maintaining sensitivity of the peripheral auditory structures and allowing external auditory input to be perceived during stridulation (Poulet & Hedwig, 2001). Through the utilization of intercellular recordings, Poulet and Hedwig (2003a, 2006) demonstrated that this central inhibition process is regulated by a specialized auditory interneuron, referred to as a "corollary discharge interneuron" which branches throughout the cricket central nervous system and has been "physiologically identified" in the prothoracic ganglion and the mesothoracic ganglion (Poulet & Hedwig, 2006). Poulet and Hedwig (2006) suggest the corollary discharge interneuron is a "multisegmental interneuron [that] is responsible for the pre- and postsynaptic inhibition of auditory neurons in singing crickets (*Gryllus bimaculatus,* p. 518)." When the motor command to move the forewings is produced, simultaneous activation of the corollary discharge interneuron occurs and Omega 1 neurons, located within the prothoracic ganglion, are inhibited. This inhibition prevents desensitization of the cricket's auditory system and allows the cricket to maintain audibility for incoming (i.e., externally produced) auditory signals during stridulation.

The echolocating bat is another animal that is highly reliant on the ability to distinguish between internal and external vocalizations. Many echolocating bats use comparisons between intense, self-generated, high frequency vocalizations and resulting reverberations or echoes to survey their surroundings (Moss & Sinha, 2003; Neuweiler, 2003). Given the intensity of these self-vocalizations would greatly surpass the intensity of the returning echo responses the question can be asked: How are these returning echo responses detected? Logically, it would seem that these echoes would be masked by the production of self-vocalizations or the bat's

auditory system would be saturated as a result of the incoming auditory input generated during self-vocalizations.

Initially, Hartridge (1945) suggested that contraction of the bat's stapedius and tensor tympani muscles regulate the attenuation of the incoming auditory signals during self-vocalizations, thus, allowing the returning echoes to be detected. Originally, however, this theory was meet with skepticism. For example, Grinnell (1963) questioned if the middle ear muscles are capable of contracting and relaxing synchronously with the rapid rates of self-vocalizations. To further investigate muscle contraction, Henson (1965) recorded cochlear microphonic and stapedius-muscle action potentials in 19 Mexican free-tailed echolocating bats during overt vocalizations. Henson demonstrated that the stapedius muscle contracted 4 to10 ms before the vocalization began, reached the point of maximum contraction just prior to sound generation, began to relax instantly post vocalization, and sustained the process of muscle relaxation throughout the vocalization. As a result of this sustained relaxation, cochlear microphonic potentials recorded to the echo responses typically had larger amplitudes than those recorded to the overt vocalizations. In general, Henson concluded that his findings supported Hartridge; however, he questioned the contribution of the tensor tympani muscle to the overall process of self-attenuation. Suga and Jen (1975) reported similar findings. These authors also noted synchronous middle ear muscle contraction to the onset of vocalization in the *Myotis lucifugus* bat species and reported that the duration of muscle contraction and relaxation produced an attenuation effect on self-produced vocalizations while maintaining auditory acuity for the returning echo responses.

At the central level, researchers have found that auditory neurons were inhibited in response to echolocating bat's self-vocalizations (Schuller, 1979; Metzner, 1989; 1993). Suga

and Schlegel (1972) and Suga and Shimozawa (1974) implemented a protocol which "compared the summated responses of the primary auditory and lateral lemniscal neurons to self-vocalized frequency modulated (FM) sounds with those evoked by the same sounds played back through a tape recorder" (Suga & Shimozawa, 1974, p.183). These authors found a simultaneous 15 dB reduction in the activity of the bat's lateral lemniscus nucleus upon vocalization. Based on this finding, Suga and colleagues speculated that neural attenuation at the level of the lateral lemniscus nucleus also along with the activation of the acoustic reflex occurs synchronously to self-vocalizations. With respect to corollary discharge, it has been suggested that in the echolocating bat, corollary discharges function to prime neurons responsive to the detection of echo responses and open temporal windows that allow for the analysis of these responses (Neuweiler, 2003; Schuller, 1979).

In more recent studies, inhibition of the bat inferior colliculus has been shown to aid in the selectivity of the bat's auditory system. Koch and Grothe (2000) investigated inhibition of the inferior colliculus in the Big Brown Bat (*Eptesicus fuscus*) and found that "neuronal inhibition sharpens tuning to the modulation frequency in the majority of neurons" (p. 71). Casseday, Covey, and Grothe (1997) showed a "class of neurons" within the inferior colliculus that responded to only sinusoidally frequency-modulated inputs, which are similar the self-vocalizations produced by the echolocating bat. Yue, Casseday, and Covey (2007) further investigated inferior colliculus inhibition in the big brown bat and found 47 of 214 neurons within the inferior colliculus responded to sinusoidally frequency modulated signals, but did not respond to other auditory inputs such as pure tones, noise bursts, single FM modulated sweeps, or sinusoidally amplitude-modulated tones.

The existence of corollary discharge systems has been described in the nonhuman primate oculomotor system (Guthrie, Porter, & Sparks, 1983) and the auditory system (Müller-Preuss, 1978; Müller-Preuss, Newman, & Jürgens, 1980; Müller-Preuss & Ploog, 1981). It has been proposed that certain cells within the primate anterior cingulate gyrus and the superior temporal gyrus are inhibited during self-vocalizations as a means to distinguish between internally generated vocalizations and externally generated vocalizations (Frith, 1992). To investigate auditory cortical differences elicited during production and perception, Müller-Preuss (1978) recorded responses from single cells within the auditory cortex of the squirrel monkey under various listening conditions (i.e., click sounds, vocalizations from other squirrel monkeys, and recorded self-generated vocalizations) presented via a loudspeaker and during self-produced vocalizations. He reported that during the listening conditions, cells within the auditory cortex produced stronger responses than when recorded during the active vocalization conditions. These initial findings were corroborated in subsequent investigations (e.g., Müller-Preuss et al., 1980; Müller-Preuss & Ploog, 1981).

Corollary discharges have been theorized to play an essential role in the learning and tuning of a songbird's mature song. Songbirds develop their mature song through a process of learning the song from a tutor (i.e., an adult bird with a developed mature song; Konishi, 2004; Seki & Okanoya, 2008) and progressing through stages where productions are not structured or systematic, to producing fewer mistakes and eventually learning the song. This process has been compared to that of human speech and language development (Doupe & Kuhl, 1999). The zebra finch is one such example. The term crystallization is used to refer to a bird's song once it is completely developed (i.e., the bird's developed song is referred as being crystallized; Leonardo & Konishi, 1999). Once the mature song is learned, it is considered stable and does not exhibit

gross acoustical variations (Brenowitz, Margoliash, & Nordeen, 1997; Konishi, 1965a, 1965b, Kroodsma & Konishi, 1991).

The underlying mechanisms contributing to and maintaining the crystallized song are unclear. Nottebohm (1968) suggested that auditory feedback was not important in maintaining the crystallized song. Conversely, Konishi (1965a, 1965b) and Scharff and Nottebohm (2001) demonstrated disruptions in song development if deafening occurred after the memorization phase of song learning, but before the song was crystallized. In additional reports, Nordeen and Nordeen (1992), Okanoya (1991), and Woolley and Rubel (1997) showed variations of the crystallized song once auditory feedback was removed (i.e., the bird was deafened). The zebra finch has also been shown to exhibit stuttering like behaviors when auditory feedback is disrupted (Cooper & Goller, 2004; Iyengar & Bottjer, 2002; Leonardo & Konishi, 1999; Mooney & Prather, 2005; Shea & Margoliash, 2003). These findings, therefore, may suggest that auditory feedback is important in maintaining the bird's song throughout life (Nordeen & Nordeen, 1992; Okanoya & Yamaguchi, 1997; Woolley & Rubel, 1997).

To further investigate how auditory feedback affected the zebra finches' song, Leonardo and Konishi (1999) monitored song changes while implementing two altered auditory feedback (AAF) conditions. The authors presented AAF to three zebra finches using an adaptive feedback perturbation paradigm and two zebra finches using a syllable-triggered perturbation paradigm. During the adaptive feedback paradigm, the zebra finches' songs were recorded and played back at a 100 ms delay, during either a silent period or in conjunction with the start of a new syllable. During the syllable-triggered perturbation paradigm a computer was set-up to recognize a target syllable. Once the target syllable was recognized a recording of that syllable was played back so that it would overlap the triggering syllable and part of the following syllable. Leonardo and

Konishi reported that both conditions of AAF caused a stuttering like disruption in the zebra finches' crystallized songs. During the adaptive feedback condition, disruptions included "stuttering" during the production of complex syllables and introductory notes. Leonardo and Konishi also noted an increase in song length along with the addition and deletion of syllables when feedback was altered. Interestingly, once the AAF was removed the bird's crystallized song was restored. Based on these results, Leonardo and Konishi concluded that auditory feedback is necessary for maintaining the adult zebra finches' crystallized song.

Altered auditory feedback and auditory monitoring. Researchers investigating the effect of AAF in individuals who stutter (IWS) and individuals who are fluent (IWF; Corey & Cuddapah, 2008) have provided support for the importance of auditory feedback during speech monitoring. Auditory feedback can be altered in both the temporal and spectral domains. For example, by delaying the time the individual hears his produced speech, delayed auditory feedback (DAF), through shifting the frequency of one's auditory input, frequency altered feedback (FAF), or through introducing a masking noise, masked altered feedback (MAF). In this review focus is given to the most commonly used methods of AAF forms, namely DAF and FAF. Experimentally delaying or shifting auditory feedback can be accomplished by having the participant speak into a microphone that is routed to an auditory input system such as a digital signal processor (DSP). The DSP then delays or shifts the transmitted speech signal from the microphone before it is presented to the listener via supraaural or insert earphones**.**

In 1950, Bernard Lee first explored DAF and made remarkable strides in the realm of stuttering research. Lee used a magnetic tape recorder and earphones to delay auditory feedback of fluent speakers by 0.04, 0.14, and 0.28 s. He noted that fluent speakers produce dysfluent speech (i.e., repetitions) when the auditory feedback was delayed. He also noted that participants

began speaking at a slower rate when auditory feedback was delayed. Lee hypothesized that delaying auditory feedback did two things to the speech of fluent speakers: disrupted natural auditory self-monitoring and slowed down speaking rate by acting as a "governor of speech rate".

These findings fueled others to investigate the effect of altering auditory feedback. Webster, Schumacher, and Lubker (1970) examined the rate of dysfluencies in six IWS while delaying auditory feedback by 100, 200, 300, 400, and 500 ms. Webster and colleagues found that the majority of their participants experienced a greater reduction in stuttering when DAF was presented at shorter delays (i.e., < 200 ms). Burke (1975) asked 20 participants who stuttered to describe pictures while listening to DAF. The DAF was varied in 50 ms steps from 0 to 300 ms. Burke concluded that the greatest reduction in stuttering was seen when DAF was also less than 200 ms (i.e., 50 to 150 ms).

Kalinowski and colleagues were the first to oppose the theory of slowed speech and suggest that the effect of altered feedback is not dependant on the rate of speech production (i.e., a reduction in stuttering episodes will occur at a slow or fast speaking rate). In their seminal study, an investigation of speaking rate and AFF on stuttering reduction was conducted (Kalinowski, Armson, Roland-Mieszkowski, Stuart, & Gracco, 1993). Nine stuttering participants read eight, 300 syllable, passages while three conditions of altered feedback (i.e., MAF, DAF, and FAF) and a non-altered feedback (NAF) were presented. The participants were instructed to read the passages at a normal speaking rate or a fast speaking rate. Stuttered episodes were counted during each condition.

In comparison to the NAF condition, a significant reduction in stuttering was demonstrated using DAF and FAF while speaking at a normal and fast speech rate. A significant

reduction in stuttering compared to NAF was not demonstrated using MAF. Kalinowski et al. (1993) concluded that a reduced speech rate was not causal of stuttering reduction when under conditions of AAF. These authors speculated that stuttering reduction under DAF and FAF is due auditory functioning. In that, during DAF and FAF conditions the participant hears their own voice altered in terms of frequency or temporal domains. IWS use this altered speech signal to increase fluency much like choral and shadowed speech is used to increase fluency. When MAF is presented the listener hears two auditory stimuli (i.e., the masker and the speaker's voice). The additional making noise possibly disrupts listener's ability to hear his own voice and again at a fast speech rate.

In subsequent studies Kalinowski and colleagues provided further empirical support against the slowed speech theory. Hargrave, Kalinowski, Stuart, Armson, and Jones (1994) examined FAF in relation to stuttering inhibition during normal and fast speech rates. The number of stuttering episodes was calculated for 14 total participants (12 males, 2 females) while reading 10 passages under NAF and four conditions of FAF (i.e., up on half octave, up one octave, down one half octave, down one octave) while speaking at normal and fast rates. Hargrave et al. reported that stuttering was equally reduced under all four FAF conditions.

Kalinowski, Stuart, Sark, and Armson (1996) found dramatic reductions in stuttering under DAF at normal and fast speaking rates. In this study, 14 adult participants, read eight passages at both a normal and fast rate while listening under four various conditions (i.e., one NAF condition and three DAF conditions). The secondary purpose of this study was to investigate how different auditory feedback delays affect dysfluency. Such as, the question was asked would a longer or shorter delay increase or decrease the total number of dysfluencies. Therefore, the DAF conditions consisted of 25, 50, and 75 ms delays. In addition to

demonstrating a reduction in stuttering at both normal and fast speaking rates, these authors found that a 50 ms delay was the shortest delay in which a reduction of stuttering was produced.

Stuart, Kalinowski, Rastatter, and Lynch (2002) investigated the effects of DAF on IWF to determine if speaking rate (i.e., a fast speaking rate vs. a normal speaking rate) influenced error production under various DAF (i.e., 0, 25, 50, or 200 ms) conditions. Sixteen fluent male participants were asked to read passages aloud into a microphone at the too speech rates, as normal (i.e., a normal rate and a normal vocal intensity) or as fast as possible. As the participants read, their speech was presented via insert earphones under various delayed conditions. There were three significant findings. It was noted that participants produced significantly more dysfluencies when reading under the DAF 200 ms condition than the other DAF conditions; significantly more dysfluencies were produced while reading at a fast rate than at a normal rate; and as the duration of the introduced delay was increased (e.g., 50 ms and 200 ms) participant's speech rate was decreased. Based on these findings it was concluded that, "temporal alterations in auditory feedback signal impact the speech-motor control system differently for people who stutter and those that do not" (p. 2239). When DAF is presented to IWS a decrease in the number of dysfluencies is evident, especially for DAF conditions of 50 ms or greater (Burke, 1975; Kalinowski et al., 1996). However, it is at DAF conditions greater than 50 ms that IWF start to experience dysfluencies.

Additionally, resurgence in the notion of auditory system involvement followed the invention of an electronic in the ear device to reduce stuttering. Kalinowski, Stuart, Rastatter, and colleagues devised an in-the-ear device referred to as the SpeechEasy™, which delivers real time DAF and FAF alterations to the speaker. This device allows IWS to benefit from the inhibitory effect of DAF and FAF without wearing bulking components such as external

microphones or body worn devices (Stuart, Xia, Jiang, Jiang, Kalinowski, & Rastatter, 2003). The SpeechEasy™ also provides a method of stuttering reduction without negating the need to maintain speech naturalness, as do most stuttering therapies. The majority of stuttering therapies focus on the employing "motoric strategies" such as slowing speech rate (Boberg, 1981) and these strategies typically result in unnaturally sounding speech and do not provide long-term results (Runyan & Adams, 1979; Runyan, Bell, & Prosek, 1990).

Researchers have also reported that fluent male and female speakers respond differently to the presentation of DAF. Sutton, Roehrig, and Kramer (1963), Bachrach (1964), and Fukawa, Yoshioka, Ozawa, and Yoshida (1988) all reported that fluent males are affected more than fluent females (i.e., produce more dysfluencies) when speaking under conditions of DAF. Correy and Cuddapah (2008) also concluded that fluent males are more susceptible to the effects of DAF than fluent females. In their study, 20 male and 21 female adult participants were asked to perform reading and conversational tasks while listening to DAF conditions of 180 ms, 230 ms, 280 ms, 330 ms, and 380 ms. Reportedly, male participants produced significantly more dysfluencies than female participants and these dysfluencies were more prevalent during the conversational tasks than the reading tasks. Corey and Cuddapah (2008) also reported that introducing DAF significantly decreased speech rate in both groups and this effect was greater during the reading tasks compared to the conversational tasks. In light of these findings, Correy and Cuddapah questioned if differences between the prevalence of male and female adults who stutter could be related to the gender differences demonstrated under DAF conditions. However, these authors noted that more research would be necessary to solidify any such relationship.

Howell, El-Yaniv, and Powell (1987) were the first of investigate the reduction of stuttering with FAF. They examined both DAF and FAF with six IWS. They reported that both

DAF and FAF conditions significantly reduced the average stuttering episodes compared to nonaltered feedback (NAF). Howell and colleagues noted that stuttering was reduced significantly more under the FAF condition compared to the DAF condition. They reasoned that this significant reduction in stuttering was demonstrated because FAF occurs at the initial onset of speech, where the majority of speech disruptions occur (Sheehan, 1974; Taylor, 1966).

Further, a disruption in speech has been found in fluent speakers when FAF was presented. Elman (1981) examined the effect of shifting the frequency of auditory feedback on five IWF. He employed two experimental conditions. During the first condition each participant was asked to maintain the production of a steady vowel token. As the participant produced this vowel, the auditory feedback was shifted from non-altered to frequency altered up or down. The second experimental condition consisted of asking each participant to repeat the phrase "Where were you a year ago?" ten times. While repeating the phrase, the participant was instructed to use an exaggerated intonation pattern and also try to say each utterance the same. Elman found that both of these conditions effected speech production. When the auditory feedback was frequency shifted up or down the participants tried to compensate for this by raising or lowering the frequency at which they spoke. Similar findings have subsequently been reported (e.g., Burnett, Freeland, Larson, & Hain, 1998; Chen, Liu, Xu, & Larson, 2007; Howell, 1990; Larson, Burnett, Kiran, & Hain, 2000; Xu, Larson, Bauer, & Hain, 2004).

Peripheral Auditory Assessment

Before continuing, a detailed review of the auditory system and clinical auditory assessments is warranted in order to provide a background for understanding the tests that were utilized in the present study. Decoding of auditory stimuli essentially begins in the auditory periphery. The peripheral auditory system has three primary functions: capturing the sound,

acting as an impedance bridge, and converting the initial mechanical signal to an electric signal, which stimulates the auditory nerve. This stimulation, in turn, begins the relay of the auditory signal through the central pathway. These functions are predominantly the result of actions performed by the outer, middle, and inner ear (Salvi, McFadden, & Wang, 2000). The basic clinical goal of peripheral auditory system assessment is to ensure the integrity of these peripheral structures. Various standardized procedures have been established to measure hearing thresholds, the ability to understand speech, and middle ear function.

Pure Tone Audiometry

The American Speech-Language Hearing Association (ASHA) defines pure tone audiometry as "the measurement of an individual's hearing sensitivity for calibrated pure tones" (2005, p. 1). These measurements or thresholds can be determined by employing manual audiometry, automatic audiometry, or computerized audiometry. This review will focus on determining hearing thresholds using manual audiometry.

According to the ASHA 2005 guidelines, pure tone thresholds should be measured using the following steps. The individual should be seated in a comfortable chair placed in a quiet or sound treated room. The examiner begins by explaining the procedure to the participant and ensures that the participant does not have any questions. The examiner also explains to the participant how they should respond when the stimulus is heard (i.e., by raising their hand or pressing a button). Once understanding of the procedure is established, the earphones are fit properly depending on the type of transducer. The examiner then presents a 1000 Hz pure tone at 30 dB hearing level (HL). If the participant hears the tone the intensity should be decreased in 10 dB steps until it is inaudible. Once inaudible the intensity is increased in 5 dB steps until the participant responds that the signal was heard. This procedure is repeated until the participant

indicates hearing the pure tone at the same intensity for two out of three trials. This is considered the hearing threshold (American National Standards Institute, 2004). The procedure is referred to as bracketing and should be completed at 250, 500, 2000, 4000, and 8000 Hz.

If the participant does not hear the 1000 Hz tone at 30 dB HL the intensity should be raised to 50 dB HL, presented again, if heard at 50 dB HL, the audiologist should continue with the bracketing procedure. If the participant does not hear the 50 dB HL tone the audiologist should increase the stimuli in 10 dB steps until the individual hears the tone.

When this procedure is conducted to determine air conduction thresholds the pure tone stimuli should be presented via insert earphones or supra-aural earphone. Air conduction thresholds referred to the thresholds obtained when the stimuli is presented along the air conduction pathway (i.e., presented through the auditory canal, middle ear, and then the cochlea). This bracketing procedure should also be used during pure tone bone conduction testing. Bone conduction testing is used to determine the hearing thresholds when the signal is presented directly to the cochlea via a bone transducer.

Pure tone testing is useful in peripheral auditory system evaluation for many reasons. Foremost, it allows for frequency specific and ear specific information that is essential for diagnostic evaluations (i.e., normal vs. abnormal hearing sensitivity), determining the etiology of hearing loss if present, and rehabilitative procedures such as hearing aid fittings. There are various degrees of hearing, ranging from normal to profound which coincide with the pure tone audiometry thresholds. Hearing thresholds are classified as follows: normal hearing is 0 to 25 dB HL, a mild hearing loss is 26 to 40 dB HL, a moderate hearing loss is 41 to 55 dB HL, a moderately-severe hearing loss is 56 to 70 dB HL, a severe hearing loss is 71 to 90 dB HL, and a profound hearing loss is 91 or greater (Goodman, 1965). These thresholds are charted on an

audiogram. The audiogram serves as a graphical representation of the individual's hearing sensitivity.

When air- and bone- conduction pure tone audiometry is used in conjunction the etiology of the hearing loss can also be evaluated. There are three types of hearing loss. A hearing loss is classified as sensorineural when the deficit is within the cochlea or along the central auditory pathway. A conductive hearing loss is when the pathology is within the outer ear, middle ear or a combination of outer and middle ear pathologies. A mixed hearing loss hearing loss refers to a hearing loss when both a sensorineural and conductive hearing loss is present. The type of hearing loss is determined through evaluating the interaction between pure tone air conduction thresholds and pure tone bone conduction thresholds. A sensorineural hearing loss is determined when the air conduction thresholds and the bone conduction thresholds are within 10 dB HL. For example, an individual has a 1000 Hz air conduction pure tone threshold of 40 dB HL and a 1000 Hz bone conduction threshold of 35 dB HL. Based on these thresholds the individual has a mild sensorineural hearing loss at 1000 Hz. When sensorineural hearing losses are found in the adult population, the common cause is typically presbycusis or noise exposure. Notably, congenital sensorineural hearing loss is documented in one in every 1000 live births (Fraser, 1964).

Conductive hearing losses manifest on the audiogram as an air-bone gap. An air-bone gap occurs when the air conduction thresholds are at least 10 dB HL poorer than the normal bone conduction thresholds. An example of this would be if the individual had a 1000 Hz air conduction threshold of 50 dB HL and a 1000 Hz bone conduction threshold of 10 dB HL. Conductive hearing losses are commonly documented in children as a result of otitis media. In adults, conductive hearing losses are typically the result of impacted cerumen. Once the cerumen

impaction is removed hearing sensitivity is restored. A mixed hearing loss is demonstrated on the audiogram as an air-bone gap (i.e., 10 dB HL difference between air and bone conduction thresholds) with both the air and bone conduction thresholds falling below normal (i.e., 25 dB HL).

Pure tone auditory testing also provides the examiner with information for determining the level other audiological evaluations should be performed (Roeser, Buckley, & Stickney, 2000). However, pure tone audiometry has been criticized as lacking validity because pure tone stimuli do not occur in everyday listening situations (Konkle & Rintelmann, 1983). Further, pure tone testing does not provide information on how the participate hears or understands spoken language. For these reasons speech audiometry is also included in the standard clinical assessment and will be reviewed.

Speech Audiometry

The peripheral auditory system can also be evaluated using speech audiometry. Speech audiometry can be helpful in that pure tone testing provides the audiologist with limited information regarding how the individual functions in everyday communicative situations. Speech measures, however, do provide insight into communicative ability and ability to perceive speech and further allow the audiologist to validate pure tone thresholds (Mendel & Danhauer, 1997). Although, there are many tests which can be used during speech audiometry speech tests which assess one's threshold for standardized speech material and assess one's speech perception for words or sentences are commonly employed (Roeser, Valente, & Hosford-Dunn, 2000). These measures typically include speech recognition threshold (SRT) testing and word recognition scores (WRS).

Speech recognition threshold. During SRT testing the examiner uses spondaic words or spondees presented to the listener via a selected transducer. Spondaic words used during testing are chosen from a list of 15 words (Young, Dudley, & Gunter, 1982). Spondaic words are two syllable words, which have equal intelligibility. They should also be familiar to the participant and representative of the English language (Hudgins, Hawkins, Karlin, & Stevens, 1947). There are several methods of determining an individual's SRT. Here the procedures recommended by ASHA and presented in the Guidelines for Determining Threshold Level for Speech (1988) are described. This procedure begins with a familiarization stage. During this stage, the examiner presents the spondaic words to the listener at a level, which is 30 to 40 dB HL above the pure tone average (PTA). The PTA is the average of threshold measures at 500, 1000, and 2000 Hz. The listener repeats the word. If correct, the examiner decreases the intensity in 10 dB steps until the listener misses two consecutive words. Once two spondees are missed the examiner increases the intensity by 10 dB HL, which is considered the starting level and actual threshold testing begins.

At the starting level, the examiner presents two words, at each level, and decreasing in 2 dB increments. Spondees are presented until the individual responds incorrectly to five out of six words presented. Testing then stops and the threshold is determined using Spearman-Karber formula (SRT = starting level - # correct + 1; Finley, 1952). If the word is repeated incorrectly the examiner the intensity is increased in 20 dB steps until the listener correctly identifies a spondee. At this point, the examiner should follow the previously described familiarization and threshold determination procedures (ASHA, 1988). As stated, SRTs are used to validate pure tone audiometry thresholds. To do so, SRTs are compared to the PTA. Brandy (2002) suggests

SRT and PTA measures are in good agreement if the are within ± 6 dB, scores differing, by 7-12 dB have are in fair agreement, and scores differing by 12 dB or more are in poor agreement.

Word recognition testing. Word recognition testing or supra-threshold speech recognition testing is used to evaluate one's ability to perceive speech at an audible level. This test can also be used to evaluate one's ability to utilize information in the frequency, intensity, and temporal domains. For example, two individuals may present similar pure tone audiograms, but their speech recognition abilities differ due to differences in the ability to use frequency, intensity, and temporal information of the speech stimuli. WRS are typically measured using phonetically balanced (PB) words presented 30 to 40 dB HL above the SRT. Speech recognition ability can be evaluated using speech materials including: nonsense syllables, digits, monosyllabic words, and sentences. This test can be administered via live or recorded voice, in quiet or in noise, with the listener typically responding by repeating the word. The examiner calculates the percentage of words repeated correctly for each ear.

As a general rule WRS are related to the degree and type of hearing loss. Those with a conductive hearing loss tend to have WRS of 90% or better, which are similar to those of a normal listener. When a sensorineural hearing loss is present the WRS will be related to the degree of hearing loss and those with retrocochlear pathologies tend to have WRS poorer than expected based on the degree of hearing loss. A common classification scheme for the determination of speech perception difficulty based on WRS is as follows: 90 to 100% = normal; 75 to 90 % = slight difficulty understanding speech; 60 to 75% = moderate difficulty understanding speech; 50 to 60% = poor recognition; scores less than 50% are considered very poor. The validity of classifying WRS in this manner has been questioned and professional agreement is lacking (Berger, 1978).

Performance intensity functions. Performance intensity (PI) functions can be utilized to evaluate the site of lesion through differentiating between cochlear and retrocochlear pathologies (Jerger, Speaks, & Trammell, 1968). Further, PI functions can be used to determine if there are dead regions along the cochlea (Halpin, 2002). PI functions are plots of WRS as a function of intensity. These functions were developed in a research effort to determine the sensation level (SL) at which the greatest intelligibility was obtained when PB words were presented. As the intensity is increased the percentage of correct words also increases. Until, a maximum score referred to as a PB-Max is reached. The PB-Max is the maximum percent of words the individual repeats correctly regardless of the intensity (Eldert & Davis, 1951).

In normal listeners, the PB-Max reaches 100% at approximately 35 to 40 dB above the SRT. In those with a sensorineural hearing loss a PB-Max of 100% will not be obtained despite the intensity level. PI functions also provide a method of identifying retrocochlear pathologies through calculating the rollover index. The rollover index is the amount at which the scores decrease once a certain intensity level is reached and is calculated by subtracting the minimum score from the maximum score (Jerger & Jerger, 1971). Rollover is commonly seen in those with retrocochlear pathologies. The degree at which rollover is considered significant varies as test characteristics vary. Jerger and Jerger (1971) presented a recorded version of the Harvard Psychoacoustics Laboratory PB-50 word lists (Egan, 1948) and determined that a rollover index of .45 was significant for classifying retrocochlear pathology. For the Northwestern University Auditory Test Number 6 (NU-6) word list, Bess (1983) concluded that a rollover index of .25 was a significant determinate of a retrocochlear pathology.

Tympanometry

The main purpose of tympanometry is to assess the middle ear function. It is important to note that tympanometry is not a measure of hearing sensitivity. There are several devices that can be used when performing tympanometry. It is important that the equipment used is capable of varying the pressure in the ear canal and producing a probe signal. Also, the equipment needed for tympanometry should have a measurement system to measure pressure changes in the ear canal.

Tympanograms are recorded by first selecting a probe tip that allows a hermetic or airtight seal to be obtained in the ear canal. The examiner then places the probe assembly on the shoulder of the participant. The pinna is lifted up and back and the probe tip placed in the ear canal until it is snug. On modern tympanometers, the examiner then presses a button to start the pressure sweep. The pressure sweep can range from + 200 daPa to – 400 daPa, which is a standard sweep or it can range from + 200 daPa to – 600 daPa considered an extended sweep. Once the pressure sweep is completed a graphical representation of the pressure changes in the ear canal will be displayed. This graph is referred to as a tympanogram. The tympanogram also provides the examiner with measurements of the tympanometric peak pressure (TPP), tympanometric width (TW), the peak compensated static acoustic admittance (Y_{tm}), and the ear canal volume (V_{ea}). From these measures the examiner is able to objectively evaluate the status of the outer and middle ear (Roser et al., 2000). According to Margolis and Hunter (2000) tympanometric values in normal adults should be as follows: Y_{tm} between 0.30 to 1.70, TW values between 51 to 114 daPa and V_{ea} between 0.9 to 2.0 cm^3. Tympanometric values falling above or below these values are indicative of a pathological condition.

Acoustic Reflex Testing

Acoustic reflex testing is another important measure of the integrity of the auditory system. It provides diagnostic information in three areas: (a) the status of the middle ear, (b) auditory nerve integrity, and (c) an indicator of auditory sensitivity (Jerger, Burney, Maudlin, & Crump, 1974). The acoustic reflex is the result of stapedial muscle contraction when a stimulus between 70 to 100 dB is presented (Northern & Gabbard, 1994). In those with normal ears this contraction causes the middle ear system to tighten or a shift in middle ear compliance. It is this middle ear tightening that is recorded during acoustic reflex testing. Clinically, an acoustic reflex threshold (ART) can be determined by measuring the lowest stimulus level at which a tightening in the middle ear is noted (Gelfand, 2002). This procedure is typically follows tympanometry. In modern equipment, the same probe assembly is used for both tympanometry and acoustic reflex testing. A 220 or 226 Hz probe tone is used expect when testing a neonate. ARTs can be measured in two ways, through ipsilateral testing and through contralateral testing. When performing ipsilateral acoustic reflex testing the probe ear and the stimulus ear are the same. In that, the probe tone is presented to the same ear which measurement is taken. When performing contralateral acoustic reflex testing the probe ear and the stimulus ear differ. In order for a contralateral acoustic reflex to be elicited the crossed pathway between the probe ear and the test has to be intact. Therefore, administrating both ipsilateral and contralateral acoustic reflex testing provides the examiner with a strong diagnostic tool (Gelfand, 2002).

In those with normal hearing and a normal acoustic reflex arc an acoustic reflex will be elicited between approximately 75 and 95 dB HL. Acoustic reflexes may be absent in those individuals with a conductive hearing loss. In the presence of cochlear pathology, acoustic

reflexes may be elevated (i.e., 100 dB HL or greater), elicited at low sensation levels (i.e., less than 60 dB HL), or absent (Wiley, Oviatt, & Block, 1987).

Loudness Balance Procedure

The loudness balance procedure is a specialized auditory assessment measure, which is useful in the separation of cochlear pathology from retrocochlear pathology. Commonly those with hearing loss will experience recruitment. Recruitment is defined by Brunt (2002) "as an abnormal growth of loudness for signals at suprathreshold intensities" (p. 111). The loudness balance procedure evaluates how a listener perceives the loudness of presented tones. The listener is presented with two tones, a constant, and a variable. He/she is asked to state if the variable tone is softer or louder than the constant tone. The results are then plotted and the pattern of recruitment determined. Typically, in those with a cochlear pathology recruitment will be seen (Hood, 1969; Hallpike, 1965; Tillman, 1969). However, recruitment will commonly not occur if the pathology is retrocochlear (Thomsen, Nyboe, Borum, Tos, & Barfoed, 1981).

Central Auditory System Assessment

Once auditory information travels through the peripheral auditory system the converted electrical signal travels along the auditory nerve entering the central auditory nervous system (CANS). The CANS is comprised of numerous decussation points where the nerve fibers crossover and comparisons from both cochleae can be made (Bhatnagar & Andy, 1995). At the level of the CANS spatial, temporal, and intensity information is gathered from the auditory signal and is represented along the auditory pathway (Phillips, 1998).

The concept of central auditory processing (CAP) is highly debated. Some researchers question the existence of central auditory processing (Rees, 1973; 1981). Those theorists apposing CAP suggest that language is a higher order function, which gives meaning to the

auditory signal (Duncan & Katz, 1983). Advocates for central auditory processing suggest that the mechanisms for processing auditory information are imperative in the learning of language (Keith, 1981). It is believed that these mechanisms may be involved in the breakdown of auditory information, which results in difficulty hearing/understanding auditory input especially in difficult listening situations. These breakdowns are referred to as a central auditory processing disorder (CAPD) and may be linked to learning, reading, and language difficulties.

There are numerous manifestations and deficit areas encompassed under the label of CAPD. ASHA (1996) characterized CAPD as one or more deficits in the following central auditory processes: sound localization and lateralization, auditory discrimination, auditory pattern recognition, and temporal processes in the absence of a peripheral hearing loss. A deficit in these areas may behaviorally manifest as difficulties with speech understanding in noise, following verbal directions, language-learning problems, and academic problems (Chermak, 2001; Musiek & Chermak, 1994). It is critical that a test battery for central auditory processing be compressive assessing the various auditory processes (Jerger & Musiek, 2000). This review will concentrate on measures implemented to evaluate how auditory information is processed through the CANS with both behavioral and electrophysiological measures.

Behavioral Auditory Processing Assessments

Dichotic listening tests. Dichotic listening tasks are composed of differing auditory stimuli presented to both ears simultaneously (Jerger & Musiek, 2000). In that, a target word is presented to one ear while a competing signal is presented to the other (Medwetsky, 2002). The purpose of dichotic listening tests are to evaluate one's ability perform binaural separation, binaural integration, and transfer between the two cerebrums via the corpus collosum. Binaural separation is the ability to hear dichotic stimuli, focus attention on a target message in one ear

and ignore the competing message in the other ear. Binaural integration conversely is the ability to hear dichotic stimuli and integrate the information presented to both ears into one message.

Anatomically, most individuals process speech in the left hemisphere. Therefore, signals presented to the right ear are lateralized directly to the left hemisphere for processing. When the target stimulus is presented in the left ear the signal must be relayed to the right temporal lobe, across the corpus callosum, and to the left temporal lobe in order for processing to occur. Kimura (1961a) suggested that the crossed auditory pathway, (i.e., when an auditory stimulus is presented to the right ear), is stronger than the uncrossed pathway (i.e., when an auditory stimulus is presented to the left ear). Kimura further suggests that the dominant temporal lobe (i.e., typically the left) is more important than the nondominate (i.e., typically the right) for the perception of spoken material.

Researchers have found an asymmetric response pattern or a cerebral dominance effect when testing normal right-handed listeners using dichotic stimuli. It was noted that during dichotic listening tasks, normal listeners tend to score higher when the target stimuli is presented in the right ear and the competing message is presented in the left ear. Compared to the target stimuli presented in the left ear while the competing signal is in the right ear (Bryden, 1963; Dirks, 1964; Kimura, 1961a; Satz, Achenback, Pattishall & Fennell, 1965). This is referred to as a right ear advantage (REA).

Based on findings from investigations administering dichotic speech tests it is suggested that these measures are possibly a diagnostic tool for not only assessing auditory function but auditory dysfunction as well (Roeser, Johns, & Price, 1976). Along with CAPD dichotic listening tests have been suggested as sensitive indictors of central nervous system lesions (Neijenhuis, Stollman, Snik, & Van den Broek, 2001), cortical and brain stem lesions (Musiek,

1983), and interhemispheric pathway lesions (B. Milner, Taylor, & Sperry, 1968; Musiek & Sachs, 1980).

Various stimuli have been used during dichotic listening tests including: tonal stimuli (Curry, 1967; Efron & Yund, 1975), sentences (Lynn & Gilroy, 1972), single words (Broadbent & Gregory, 1964), digits (Broadbent, 1954; Kimura, 1961b), and nonsense syllables (Berlin, Lowe-Bell, Cullen, Thompson, & Loovis, 1973). However, dichotic digits and dichotic consonants are commonly employed during dichotic listening tests. Dichotic digits and dichotic consonant test procedures consist of presenting two overlapping digits or consonant to both ears simultaneously. Typically the listener is asked to verbally indentify either digits or consonants. Dichotic listening ability is calculated as the percent correct. One dichotic digit test modified by Musiek (1983) scores between 100 to 90% is considered normal when normal hearing was present. If a score of below 90% was obtained in those with normal hearing the finding was considered abnormal.

Broadbent (1954) was the first to present dichotic digits using a dichotic listening paradigm. Kimura pioneered the use of dichotic listening techniques in those with cortical dysfunction. Kimura (1961b) administered the dichotic digits test to those who had under gone removal of tissue (i.e., ablations) located in the temporal lobe. She reported that those participants demonstrated lower scores in the ear contralateral to the lesion.

Synthetic sentence identification with contralateral competing message. The synthetic sentence identification with contralateral competing message (SSI-CCM) is a dichotic speech test, which evaluates one's ability to separate auditory signals presented at the same time. During this test ten third-order approximations of English sentences (Speaks & Jerger, 1965) are presented in the target ear while a competing message is continually presented in the other ear.

The target message is presented at 30 dB HL while the competing message is presented at varying levels from 30 dB HL (i.e., 0 dB signal-to-noise [S/N]) to 70 dB HL (i.e., -40 S/N). Variations in presentation level are performed in order to evaluate the listener's auditory separation ability at different S/Ns or message-to-competition ratios. The listener is given a list of the 10 sentences and is asked to state the number of the sentence heard. The SSI-CCM is then scored in terms of percent correct at the most difficult listening situation (i.e., -40 S/N) or as an average of the percent correct during each listening condition.

In terms of diagnostic utility the SSI-CCM has been suggested to be an adequate tool for differentiating between brainstem and cortical dysfunctions (Jerger & Jerger, 1974; 1975; Keith, 1977). In normal adults and adults with brainstem lesions typically score around 100% for all listening conditions. In adults with temporal lobe pathologies a reduced score is typically noted in the ear contralateral to the side of lesion (Jerger & Jerger, 1975). This test can also be administered with the competing message in the same ear (i.e., with an ipsilateral competing message) and will be discussed in a later section.

Staggered spondaic words. Staggered spondaic words (SSW) developed by Katz (1962) are a commonly used dichotic speech test. The SSW is also sensitive to brainstem and cortical pathologies (Katz, 1962). This test utilizes spondaic words presented so they overlap. In that, during SSW administration one spondee is presented to one ear so that the second syllable overlaps the presentation of the first syllable of the word presented in the other ear; resulting in the listener hearing one word (Katz & Ivey, 1994). For example, the word "stairway" is presented to the left ear while the word "farewell" is presented to the right ear. The listener will respond that the word "stairwell" was heard. Further, this overlapping allows the words to be presented in four conditions: right noncompeting (RNC), right competing (RC), left competing

(LC), and left noncompeting (LNC) and vary between right ear first presentations and left ear first presentations. The RC and LC conditions assess the listener's binaural integration ability; where as, the RNC and LNC conditions assess word recognition ability (Schow & Chermak, 1999). This test is presented at 50 dB HL above the PTA and is scored in terms of percent of errors in each, in each condition, and the total percent of errors.

Scoring is performed and response patterns determined. These response patterns are then related to auditory processing subtypes (i.e., certain patterns are consistent with certain auditory processing deficits; Katz, Smith, & Kurpita, 1992). However, the development of normative data is critical for the interpretation of this evaluation (Bellis, 2003).

Monaural low-redundancy speech tests. Monaural low-redundancy speech tests address the listener's ability to use intrinsic and extrinsic redundancy cues found in the auditory system and language structure to process a message in degraded listening situations. More specifically monaural low-redundancy speech tests evaluate one's ability to perform auditory closure. Auditory closure refers to the listener's ability to understand a message when parts of the message are missing or unclear. In normal listeners the auditory system is highly redundant allowing auditory closure to take place. Those with auditory closure deficits experience breakdowns in the auditory system's redundancy and are unable to perform auditory closure. To evaluate auditory closure, degraded auditory signals are presented so the listener must utilize the auditory system's intrinsic redundancy. Degrading the auditory signal can be accomplished by, presenting the stimuli with a competing message or noise, low-pass filtering the stimuli (Bocca, Calearo, & Cassinari, 1954), or time compressing the stimuli (Bellis, 2003).

Synthetic sentence identification with ipsilateral competing message. One monaural low-redundancy measure is the synthetic sentence identification with ipsilateral competing

message (SSI-ICM) test. During the SSI-ICM both the target stimuli and the competing message are presented to the test ear. As in the SSI-CCM procedure ten third order English sentences are presented while a competing message is heard. The primary message is presented at 30 dB HL and the competing message is presented at various intensities from 20 dB HL (i.e., +10 S/N) to 50 dB HL (i.e., -20 S/N). Again, varying the S/N allows for the evaluation of redundancy skills as the difficulty in listening situation increases. The listener is given a list of the ten sentences and responds by stating the number of the sentence heard. The SSI-ICM is scored by averaging the percent correct at 0 dB S/N and -20 dB S/N. Scores typically seen in normal adult listeners are as follows: 100% at a S/N of 0 dB, 80% at a S/N of -10 dB, 35% at a S/N of -20 dB, and 20% at a S/N of -30 dB (Jerger, 1973).

Time compressed speech tests. Other measures of auditory closure include time compressed speech tests. Time compressing the auditory stimuli involves reducing the duration of the signal at various compressions, while the frequency characteristics remain the same (Fairbanks, Everitt, & Jerger, 1954). Commonly, NU-6 words are presented under time-compressed conditions and the listener verbalizes what he/she hears. Scores are determined as percent correct in each ear. In normal listeners a reduction in score will be noted as the percent of time compression increases. These measures are suggested as sensitive indicators of diffuse primary auditory cortex pathologies especially when the stimulus is highly compressed (Baran, Verkest, Gollegly, Kibbe-Michal, Rintelmann, & Musiek, 1985; Kurdziel, Noffsinger, & Olsen, 1976; Mueller, Beck, & Sedge, 1987). In normal listeners, using NU-6 words, typical scores are 86.5% at 45% compression and 55.5% at 65% compression. Scores falling below two standard deviations below these normative values would be considered abnormal (Wilson, Preece, Salamon, Sperry, & Bornstein, 1994).

Binaural interaction tests. Binaural interaction refers to the interaction of both ears (i.e., working together) and is essential for auditory processes such as: sound localization (i.e., determining sound's direction); sound lateralization (i.e., the perception of a sound's place); masking level difference; hearing in noise; and binaural fusion.

Sound localization and lateralization. Sound localization and lateralization are important in the processing of linguistic information, in that one must realize the presence of a sound before determining its importance. For example, before further processing takes place, one must be able to localize the sound's source and importantly differentiate between the source of a target message and the source of a competing message. Normal listeners typically localize sound with 100% accuracy and need an interaural intensity difference of approximately 4 dB to lateralize sound (Pinheiro & Tobin, 1969, 1971). Peripheral pathologies such as unilateral hearing loss (Nordlund, 1964) have shown to affect sound localization and lateralization abilities. Further, investigations of the mechanisms underlying sound localization and lateralization suggests that low brainstem, auditory nerve, temporal lobe, and corpus collosum lesions also degrade one's ability to perform these functions through decreasing accuracy and increasing interaural intensity difference (Nordlund, 1964; Pinheiro & Tobin, 1969, 1971; Sanchez-Longo & Forster, 1958).

Binaural release from masking. A listener's thresholds increase when a masking noise is presented in phase (i.e., homophasic) with the target stimuli (i.e., pure tone or speech). However, when the masking noise is presented out of phase (i.e., antiphasic) with either the speech-tone or with the other ear a decrease in listener thresholds will occur (Bellis, 2003; Chermak & Musiek, 1997). The phenomenon is referred to as binaural release from masking. Simply, the listener is released from the effect of masking during binaurally presented antiphasic

conditions; resulting in an improvement in listener thresholds compared to homosphasic conditions.

In a seminal investigation, Licklider (1948) examined the binaural release from masking using bilaterally presented simultaneous speech and noise stimuli. During stimulus presentation Licklider varied the interaural phase relationship between the ears to determine how these variations affect speech intelligibility. Licklider found that normal listeners demonstrated a reduction in speech intelligibility during homophasic conditions compared to antiphasic conditions. Further, increased speech intelligibility was demonstrated during binaural antiphasic conditions compared to monaural presentations. When the speech and noise were presented in a binaural homophasic condition, speech intelligibility was reduced compared to monaural presentations. Based on these findings Licklider concluded that there is possibly a relationship between interaural inhibition and interaural phase.

In a subsequent study, Hirsh (1948) employed six experimental conditions to examine the affects on interaural phase and pure tone frequency and intensity on interaural summation and inhibition. The six conditions Hirsh used were as follows: (a) binaurally presented noise with monaurally presented tones, in phase (SNø); (b) binaurally presented noise with monaurally presented tones, 180^{o} out of phase (SNπ); (c) binaurally presented noise and tones, in phase (SøNø homophasic); (d) binaurally presented noise and tones, out of phase (SπNπ homophasic); (e) in phase binaurally presented tones with binaurally presented out of phase noise (SøNπ antiphasic); (f) out of phase binaurally presented tones with binaurally presented in phase noise (SπNø antiphasic). Hirsh found that noise presented binaurally 180^{o} out of phase was localized at the ear level; where as, homophasiclly presented noise was localized at the midline. Therefore,

monaural pure tone thresholds increased during homophasic conditions and binaural pure tone thresholds increased during antiphasic conditions (i.e., binaural release from masking).

Typically, the binaural release from masking phenomenon is evaluated in terms of the difference between binaural thresholds obtained during homophasic and antiphasic conditions. This difference is referred as the masking level difference (MLD; Hall, Buss, & Grose, 2007). MLDs can be obtained using speech or nonspeech stimuli under various experimental parameters. These parameters consequently affect MLD performance. Clinically, MLDs are commonly calculated by subtracting the listener's performance during the SπNø condition with listener performance during the SøNø condition. An MLD of 5.5 dB or greater is considered normal; where as, scores lower than 5.5 dB are considered abnormal (Wilson, Zizz, & Sperry, 1994). Lynn, Gilroy, Taylor, and Leiser (1981) concluded that abnormal MLDs are indicative of low brainstem lesions, in that, cortical pathologies and lesions occurring higher in the brainstem do not affect MLD (e.g., scores 5.5 dB or greater).

Binaural fusion. Binaural fusion (BF) tasks evaluate the ability to fuse auditory information presented to both ears into one (Bellis, 2003). Typically, during these tasks one low-pass filtered word is presented to one ear and one high-pass filtered word to the other ear at 20-30 dB SL (Willeford, 1977). The listener is asked to verbalize what is heard and performance is scored in terms of percent correct. In normative studies performed by Willeford and Burleigh (1985) using the Ivey BF task presented at 30 dB SL normal listening adults tended to score around 90%. In terms of pathological conditions, it is suggested that BF is mediated at the level of the brainstem and is greatly affected by brainstem pathologies (Lynn & Gilroy, 1975; Smith & Redneck, 1972). Metzner (1959) reported that those with brainstem pathologies demonstrated a marked reduction in BF ability compared to normal listeners***;*** where as, those with cortical

pathologies performed similarly to normal listeners.

Temporal processing tests. Time related aspects of auditory processing are referred to as temporal processing. ASHA (1996) stated that temporal processing encompasses various auditory processes such as temporal resolution, masking, temporal integration, and temporal ordering. These processes are fundamental for auditory skills including speech perception, rhythm perception, gap detection, pitch discrimination, and auditory figure ground segregation (Hirsh, 1959; Phillips, 2002).

Temporal processing abilities can be evaluated using various methods. Commonly employed, are tasks that evaluate temporal order or temporal patterning. During these tasks the listener determines either the frequency (i.e., high or low) or the duration (i.e., short or long) of tonal stimuli. In order for one to be able to sufficiently perform these tasks (i.e., perceived and verbally labeled) both left and right temporal lobes and the corpus collosum must be intact (Musiek, Pinheiro, & Wilson, 1980; Pinherio, 1976). In that, possessing the pattern (i.e., pitch) of the presented tones typically takes place in the right hemisphere (Jerger, 1997); where as, the left hemisphere controls the verbal labeling of the perceived pattern (Gazzaniga & Sperry, 1967). The corpus collosum transfers the information between the right and left hemispheres. If lesions were present within the right hemisphere, left hemisphere, or the corpus collosum the listener would demonstrate abnormal responses for these measures.

Electrophysiological Auditory System Assessment

Technological advancements have allowed researchers to investigate central laterality, anatomical structures of the auditory system, and central auditory processing abilities through measuring the electrophysiological brain activity in response to sound. These measurements are

commonly referred to as auditory evoked potentials or auditory event-related potentials (ERPs). Fortunately, ERPs allow one to evaluate the auditory system with non-invasive procedures.

Electrophysiological auditory system assessment consists of the placement of small electrodes on the face and scalp, presenting an auditory stimulus (e.g., a click, tone burst, or speech) and measuring the brain activity elicited by the auditory stimuli (Alain & Tremblay, 2007). In those with normal hearing and a normal central auditory pathway, the elicited brain activity will produce a series of positive and negative peaks, which comprises an electroencephalogram waveform or tracing. The evoked waveform is typically analyzed in terms of morphology, latency, and amplitude. ERPs are classified as early, middle, or late potentials depending on the time of in which they occurrence post stimulus onset.

To fully understand electrophysiological auditory assessment it is essential to understand how these measures are analyzed; therefore, a review of parameters commonly evaluated will be given. ERPs recorded from a normal listener will be composed of positive and negative peaks and will be analyzed in terms of morphology, latency, and amplitude. Morphology is a subjective judgment of the waveform's general shape (i.e., the examiner judges the shape of the waveform and determines if the waveform looks as would be expected). The latency of the waveform is the time in which a peak (i.e., positive or negative) occurs in relation to the stimulus onset, usually measured in milliseconds (ms). Alain and Tremblay (2007) described latency as associated with neural conduction stating that "latency is therefore related to neural conduction time and site of excitation: the time it takes the sound to travel through the peripheral auditory system to the place of excitation in the central auditory system" (p. 576). Amplitude reflects "the size of the neural population indexing processing" (Martin, Tremblay, & Korczak, 2008, p. 285) and is measured in microvolts (μV). Amplitude is typically calculated as the positive to proceeding

negative peak difference. A positive peak, denoted by a P, is a point where the where amplitude is high. A negative peak or trough, denoted by an N, is a point in the waveform where the amplitude is low. These waveforms are represented along a time domain (i.e., amplitude as a function of time). Classifications of early, middle, or late are given to the responses depending on where they the positive and negative peaks occur along the time domain.

Auditory Brainstem Responses

Auditory brainstem response (ABR) recordings are early evoked potentials that occur within 10 ms of the stimulus onset in the normal adult population (Burkard & Secor, 2002; Dyson & Alain, 2004). This response consists of five to seven waves that measure the synchrony of neural firing of the eighth nerve and auditory brainstem (Don & Kwong, 2002; Song, Banni, Russo, & Kraus, 2006). Researchers have performed many investigations looking into the neurological generators of ABR waveforms in both humans and animal models. Although, the exact neurological locations of Waves I-V are unknown, it is speculated that responses are recorded from anatomical structures ascending along the auditory pathway. Wave I is generated from the distal portion (closest to the cochlea) of the auditory nerve (Hashimoto, Ishiyama, Yoshimoto, & Nemoto, 1981; Møller, 1985; Møller & Jannetta, 1981, 1982, 1983). Wave II is also considered to originate from the eighth nerve, but from fibers closer the brainstem and cochlear nucleus (Lazorthes, LaComme, Ganbert, & Planel, 1961). Wave III is suggested to originate from the superior olivary complex (Buckwald & Hung, 1975; Lev & Sohmer, 1972), Wave IV from the mid and upper pons (Guilhoto, Quintal, & da Costa, 2003), and wave V the lateral lemniscus or the inferior coliculus (Guilhoto et al., 2003). ABR responses are used in the clinical setting to estimate of hearing sensitivity in populations where objective pure tone

measurements cannot be obtained such as in infants and those with mental handicaps. Given, that the ABR response is not susceptible to state of arousal infants can be evaluated while sleeping.

In normal adults, ABRs are recorded by placing small disc electrodes on the scalp and mastoids while an auditory stimulus is presented trough a transducer. Several auditory stimuli can be used (i.e., clicks, tone-bursts, or speech), but click stimulus is routinely used to evoke ABR responses. The click's abrupt onset and brief duration contribute to good synchronization, minimize stimulus artifact, and provide a broad spectrum that stimulates many nerve fibers. Click stimuli can be varied in terms of rate. In that, clicks can be presented at a slow or fast rate and ABR responses are affected as a consequence of this rate.

An important parameter for ABR waveform analysis is latency. Latency is measured as the time between stimulus onset and a change in the waveform. In the normal ABR five to seven prominent peaks are recorded and are labeled using Roman numerals I-VII. In all normal adults, peaks elicited with an audible stimulus should occur within approximately the same time frame and have a similar morphological appearance. The approximate times for ABR peaks in normal adults are as follows: wave I ~1.54, wave II~ 2.67, wave III~3.73, wave IV~4.81, and wave V~5.52 (Antonelli, Bellotto, & Grandori, 1987). ABRs are also commonly analyzed in terms of interpeak latencies. Interpeak latencies can be defined as the time between one peak and another peak (Hall, 1992). I-III, III-V, and I-V interpeak latencies are thought to reflect the conduction time along the auditory pathway (Elidan, Sohmer, Gafni, & Kahana, 1982; Fabiani, Sohmer, Tait, Gafni, & Kinarti, 1979; Griffiths, Chambers, & Bilger, 1989).

Amplitude measures are another important parameter used in the analysis of ABR waveforms. Again, amplitude refers to the size of the ERPs wave component expressed in microvolts (μV) and is typically measured from the peak of one component to the following

trough. For ABR waveforms amplitudes commonly do not exceed 1.00 μV. ABR amplitudes are affected by factors such as muscle artifact, stimulus intensity, and filter settings making ABR amplitudes highly variable. Researchers have suggested analyzing the wave V/I amplitude ratio instead of absolute ABR amplitudes in order to limit the variability found in this parameter (Chiappa, Gladstone, & Young, 1979; Rowe, 1978; Starr & Achor, 1975). Although, these responses are also variable, a wave V/I ratio less than 1.00 μV but greater than 0.5 μV is demonstrated in approximately 10% of the population. Starr and Achor (1975) concluded that a wave V/I ratio of 0.5 μV is at the lower normative trend.

Pathological conditions affect ABR waveforms in various ways. In those with a conductive hearing loss a common findings is a prolongation of the wave V latency. This prolongation is also recorded in those with retrocochlear pathologies. Recordings from an individual with a sensorineural hearing loss will also produce a wave V latency prolongation. When the waveform is elicited using a click stimuli the degree of prolongation will be influenced by the degree of hearing loss at 4000 Hz (Coats, 1978; Coats & Martin, 1977; Jerger & Mauldin, 1978). Therefore, evaluation of all waveform components as well as the stimulus parameters used to evoke the response is important for a sound diagnosis.

Auditory Middle Latency Responses

The auditory middle latency response (MLR) is an auditory evoked response occurring between 12-50 ms (Hall, 1992) or 10-80ms (Kraus & McGee, 1993) post stimulus onset. These responses are classified as a series of four waves two negative (Na, Nb) and two positive (Pa, Pb). Clinically MLRs are used to assess low frequency hearing loss, auditory pathway site of lesion, auditory functioning, and cochlear implant functioning (Kraus & McGree, 1993). However, clinical usefulness of the MLR is limited. In that, MLRs are affected by sleep state and

sedation and are inconsistently recorded from infants. Kraus, Smith, Reed, Stein, and Cartee (1985) reported that only 20% of normal infants demonstrated Pa responses.

The neurological origins of these responses have not been solidified; however, current animal studies have provided insight into the neurological generators. Na has been suggested to originate from deep sub-cortical regions such as the inferior colliculus and the medial geniculate body of the thalamus. The Pa component is suggested to originate from the auditory cortex located in the posterior temporal lobe (Deiber, Ibanez, Fischer, Perrin, & Mauguiere, 1988; Jacobson & Newman, 1990; Lee, Lueders, Dinner, Lesser, Hahn, & Klem, 1984). The Pb component is also suggested to originate from temporal lobe but from regions, which are more anterior and lateral to the Pa component (Scherg, Hari, & Hämäläinen 1989; Yoshiura, Ueno, Iramina, & Masuda, 1995). Additionally, it has been suggested that MLRs are generated from the interaction between the primary pathways including the "ventral portion of the medical geniculate body and the primary auditory cortex" (Kraus & McGee, 1993, p.44) and the nonprimary pathways. Kraus and McGee (1993, 1995) noted that the primary pathways are sensitive to auditory stimuli, display frequency tuning, time-locking abilities, binaural interactions where as the non-primary pathways have multiple functions, broadly tuned, and demonstrate mild binaural interactions.

A standard electrode montage for recording an MLR has not been established. MLR recordings can be elicited using a single or multiple channel recording montages, linear arrays, or incorporated into measurements evoking more complex brain images. When recorded in normal adults, at frontal midline electrode sites (Fz and Cz), using click stimuli, MLR waveform components will typically demonstrate the following latency values: Na (latency: ~19 ms), Pa (~30 ms), Nb (~40 ms), and Pb (~50-70 ms; Yvert, Crouzeix, Bertrand, Seither-Preisler, &

Pantev, 2001). MLR waveforms are analyzed differently than other evoked potentials, in that, latency and amplitude measures are not crucial factors in determining if the response is normal or abnormal. Normal MLR waveforms may not contain all components (i.e., Na, Pa, Nb, and Pb). The Pb component is an unpredictable response and is often absent in normal listeners. Therefore, normalcy of the MLR response is typically based on the presence of Na and Pa waveform components, their morphology, and repeatability.

Pathological conditions have shown to variably affect MLR recordings. In terms of hearing assessments, MLRs can be used as an assessment of conductive, sensorineural, and mixed hearing losses (McFarland, Vivion, & Goldstein, 1977). If the response is elicited with tonal stimuli from an individual with a low frequency hearing loss, recordings should reflect the degree of hearing loss. Researchers have also found interesting results when employing MLRs in those with cortical pathologies. Kileny, Paccioretti, and Wilson (1987) reported that MLRs are absent or have reduced amplitudes when recorded from individuals with a lesion in the primary auditory cortex. However, those with lesions in the medial geniculate body, auditory association areas, or frontal or parietal operculum areas MLRs were not affected.

Auditory Long Latency Event-Related Responses

Auditory long latency ERPs are a series of event-related potentials occurring between 50 to 500 ms after stimulus onset (Goff, 1978; Hall, 1992). Long latency ERPs are comprised of a series of positive and negative peaks labeled P1-N2-P2 along with the mismatch negativity (MMN) and the P300 response. These potentials are classified as either endogenous or exogenous responses. Exdogenous potentials are elicited in response to the auditory stimuli (Goff, 1978). The grouping of exdogenous responses comprises most auditory evoked potentials (e.g., ABR, MLR, and P1-N1-P2 complex). Goff (1978) stated that exdogenous responses “are

considered to be principally determined by the evoking stimulus parameters and to reflect "obligatory" neural events" (p. 519).

The MMN and P300 are endogenous responses. Endogenous responses are evoked by the listener's cognitive reaction to the auditory stimulus. Controversy still surrounds the neurological generation sites of long latency ERP components; however, several generator sites have been suggested including the temporal lobe (Goff, 1978), supra-temporal plane and the auditory cortex (Hari, Aittoniemi, Jarvinen, Katila, & Varpula, 1980; Näätänen & Picton, 1987; Scherg & von Cramon, 1985; Wood & Wolpaw, 1982) and the primary and association auditory cortices (Fitzgerald & Picton, 1983; Ballantyne, 1990).

Mismatch negativity. The MMN is an automatic endogenous response that represents listeners' reactions to an unexpected change in the auditory stimulus (Kraus, McGee, Carrell, & Sharma, 1995; Näätänen, Pakarinen, Rinne, & Takegata, 2004; Näätänen, 1992) and is not affected by listener attention (Näätänen, 1990; Sussman, Winkler, & Wang, 2003). The response occurs at approximately 100 to 200 ms, is a negative deflection, and is evoked by way of an oddball paradigm. An oddball paradigm consists of a series of oddball or "target" stimuli presented pseudo-randomly within a series of standard stimuli. MMN is thought to reflect a preattentive perceptual response that allows attention to be switched between auditory stimuli presentations.

Näätänen, Gaillard, and Mantysalo (1978) hypothesized that the MMN results from a mismatch between the standard stimuli and the deviant stimuli. They proposed that the standard stimulus formulates a cognitive memory trace in which the unexpected auditory deviant is compared. Therefore, the MMN is generated as a response to the changing auditory event and could be used as a measure of auditory discrimination.

The MMN originates from multiple neurological generators, with the supratemporal auditory cortex being the major neurological generator. Other generator sites include the frontal and subcortical regions (Alho, Sams, Paavilainen & Näätänen, 1986; Näätänen, 1990). Latency and amplitudes parameters are used in the analysis of MMN waveforms. These measures are thought to reflect the listener's ability to accurately perceive auditory stimuli and also reflect the degree to which the standard and target stimuli differ. When stimulus characteristics (e.g., frequency or duration) between the standard and the deviant become increasing different, MMN latency will decrease and MMN amplitude will increase. In pathological populations a greater increase will be needed for the expected latency–amplitude shift to occur. Therefore, MMN has been suggested as a clinical tool for the assessment of auditory discrimination.

P300. The P300 is a robust auditory late evoked potential, discovered in 1965 by Sutton, Braren, Zubin, and John. The P300 is thought to represent a neural index of higher level cortical functioning such as information processing and memory (Donchin, 1981; Squires, Donchin, Squires, & Grossberg, 1977; Harrison, Buchwald, & Kaga, 1986; Ford, Mohs, Pfefferbaum & Kopell, 1980; Squires, Wichens, Squires, & Donchin, 1976). This potential is typically elicited by changes in the auditory stimuli (Goff, 1978; Morgan, Cranford, and Burk, 1997) or an oddball paradigm between 250 and 500 ms. Two waves are elicited when recording the P300, P3a and P3b. The P3a is independent of listener's attention and is evoked when the standard and the deviant largely differ. Research has suggested that the P3b response is representative of the listener's ability to discriminate the changes in the auditory stimuli; therefore, this response is affected by attention (Morgan et al., 1997) in that the listener must be conscious to the changes in the auditory stimuli (Polich, 1989; Squires, Squires, & Hillyard, 1975).

To facilitate listener attention, he/she is typically asked to perform a task while the response is recorded. This task may be a mental activity, such as, counting the number of presented deviant stimuli (Courchesne, 1978; Kurtzberg, Vaughan, & Kreuzer, 1979) or may be a psychical activity, such as, pressing a response button when a deviant stimulus is heard. The type of activity used does not affect P300 activation nor does the response activity affect latency or amplitude measurements (Kileny & Kripal, 1987; Sklare & Lynn, 1984).

Several researchers have theorized as to what mechanisms underlie the P300. Desmedt (1980) hypothesized that the occurrence of the P300 signals the end of recognition or the end of a processing period. Donchin and Coles (1988) suggested that P300 is manifested as a result of contextual updating in the working memory. Neurologically, several cortical and sub-cortical sites have been implied as P300 generators. Suggested cortical generator sites include frontal regions (Courschesne, 1978; Desmedt & Debecker, 1979; Wood & McCarthy, 1986), centro-parietal regions (Goff, Allison, & Vaughan, 1978; Pritchard, 1981; Simson, Vaughan, & Ritter, 1977); the temporal parietal junction (Knight, Scabini, Woods, & Clayworth, 1989); and auditory cortical regions (Richer, Johnson, & Beatty, 1983; Rogers, Baumann, Papanicolaou, Bourbon, Alagarsamy, & Eisenberg, 1991). Subcortical generator sites include the hippocampus (Halgren, Squires, Wilson, Rohrbaugh, Babb, & Crandall, 1980; Okada, Kaufman, & Williamson, 1983; Squires, Halgren, Wilson, & Crandall, 1983) and the thalamus (Wood, Allison, Goff, Williamson, & Spencer, 1980). Further, neural pathways such as the mesencephalic reticular formation, medial thalamus, and the prefrontal cortex have been implied as P300 generator sites because attention is essential for the response to be elicited (Yingling & Hosobuchi, 1984; Yingling & Skinner, 1977).

For P300 recordings electrodes are commonly placed along the midline of the scalp where the maximum response is recorded (Polich & Starr, 1983). P300 response waveforms are analyzed in terms of latency and amplitude. The latency of the P300 response is considered a reflection of stimulus evaluation (Kutas, McCarthy, & Donchin, 1977) or the speed in which neural processes (i.e., attention allocation and memory) occur (Duncan-Johnson, 1981; Magliero, Bashore, Coles, & Donchin, 1984). If the oddball paradigm used to elicit the P300 consists of standard and deviant stimuli are difficult to discriminate (i.e. the processing demand is great) the P300 latency will increase (Naylor, Halliday, Callaway, Yano, & Walton, 1987). Amplitude of the P300 response is considered to be a reflection of attention (Sommer, Matt, & Leuthold, 1990). In normal adult listeners, P300 amplitudes increase as the standard and deviant become easier to discriminate. For example, if the P300 is elicited with standard and deviant stimuli that are easily discriminated the P300 latency will decrease and the P300 amplitude will increase. If standard and deviant discrimination difficulty is increased the P300 latency will increase and the P300 amplitude will decrease.

How does neurological pathology affect the P300 response? This question has been explored in numerous clinical populations; however, due to intersubject variability, determining pathological effect on the P300 response has proven to be difficult. Picton and Hillyard (1988) suggested using the P300 response in any disorder as a correlative of global cognitive dysfunction. Goodin, Squires, and Starr (1983) and Polich and Starr (1983) suggested using the P300 response as a monitor of therapy effectiveness, given that, intrasubject recordings do not vary greatly and P300 latency reflects an individual's cognitive processing demand. If the P300 response were recorded before and during treatment using the same recorded parameters an objective measure of treatment effectiveness would be obtained. For example, if the P300

latency decreases while the listener is undergoing treatment (i.e., demonstrated an increase in processing ability) the examiner could assume that the treatment is effective. Conversely, the examiner would know to change the treatment method if the P300 latency does not decrease (i.e., demonstrating processing ability is not improving).

P1-N1-P2 Auditory Event-Related Potential

P1-N1-P2 waveform components are auditory long latency ERPs occurring between 80 and 250 ms (Martin & Stapells, 2005; Sharma, Marsh, & Dorman, 2000). These potentials were first discovered in 1939 by Pauline Davis and were the first auditory evoked potentials to be recorded (Hall, 1992). It was not until the early 1960s, with the advancements in averaging and filtering techniques that these recording potentials grew in popularity and implementation. Auditory ERPs have been explored across a vast array of clinical populations and have subsequently become accepted as a reliable objective measure of the auditory system (Hall, 1992).

The P1-N1-P2 complex is considered to be an obligatory response, meaning that waveform components are essentially controlled by stimulus characteristics and not the perception of the auditory cues by the listener. This response is thought to reflect speech representation in the central auditory system independently of listener attention (Martin, Sigal, Krutzberg, & Stapells, 1997; Ostroff, Martin, & Boothroyd, 1998; Sharma & Dorman, 1999, 2000; Whiting, Martin, & Stapells, 1998). In addition, the N1-P2 complex has been suggested to be a representation of the sensory encoding of auditory stimulus characteristics (Näätänen & Picton, 1987). The N1 component is considered to represent the onset of the response. In other words, the N1 proportion of the P1-N1-P2 complex is generated approximately 100 ms following the beginning (i.e., onset) of an auditory stimulus **(**Alain & Tremblay, 2007).

Neurophysiological Origin

The P1-N1-P2 complex is generally referred to and investigated as one unit (Crowley & Colrain, 2004); however, researchers have suggested that each component may represent synchronous activity from separate cortical sources. The primary auditory cortex has been suggested as the generator for the P1 component (Çelik, Seleker, Sucu, & Forta, 2000; Eggermont & Ponton, 2003; Liégeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994; Woods, Clayworth, Knight, Simpson, & Naeser, 1987); however, studies investigating this component are sparse, as researchers have mainly focused on identifying the generators of the N1-P2 components.

The N1 component is believed to represent synchronous neural activity from multiple generators within the primary and secondary auditory cortex (Martin et al., 2008; Näätänen, 1992; Näätänen & Picton, 1987) and thalamic-cortical radiations to the presentation of an auditory stimulus (Näätänen & Picton, 1987; Picton et al., 1999; Woods, 1995; Wolpaw & Penry, 1975). Other researchers have suggested a supratemporal neural generator (Hari et al., 1980; Scherg & von Cramon, 1985; Wood & Wolpaw, 1982). There are three main components of the N1 response, N1b, the "T-Complex", and a negative deflection at approximately 100 ms, which have differing characteristics and are maximally generated from differing electrode sites. These components are speculated to originate from differing anatomical generators.

In most studies the N1 is described in relation to the corresponding P2 component; however, as stated above, the P2 may represent activity from different neural generators than the N1 (Colrain & Crowley, 2004; Roth, Ford, Lewis, & Kopell, 1976; Vaughan, Ritter, & Simon, 1980). Initially, the temporal lobe, more specifically the auditory cortex, was considered as the generating source of the P2 component (Elberling, Bak, Kofoed, Lebech, & Saermark, 1980;

Hari, et al., 1980; Perrault & Picton, 1984). More recently, localization and imaging studies have been used to suggest multiple generators of the P2 component (Jacobson, 1994). These speculated generators include the mesencephalic reticular activating system (Knight, Hillyard, Woods, & Neville, 1980), a component of the reticular formation in the brainstem and is responsible for transmitting sensory input (Beine, 2007), the planum temporale and the auditory association complex Area 22 (Godey, Schwartz, de Graaf, Chauvel, & Liegeois-Chauvel, 2001), and the Sylvian fissure at the level of the somatosensory area S2 (Hari, Hamalainen, Kekoni, Sams, & Tiihonen, 1990). Others have shown that the P2 is generated from sources located anterior to that of N1 generators (Hari, Pelizzone, Makela, Hallstrom, Leinonen, & Lounasmaa, 1987; Pantev, Hoke, Lütkenhöner, & Lehnertz, 1991, Pantev, Eulitz, Hampson, Ross, & Roberts, 1996; Rif, Hari, Hamalainen, & Sams, 1991; Sams, Paavilainen, Alho, & Näätänen 1985). Through the utilization of magnetoencephalograhy (MEG), Hari et al. (1987) suggested that generated wave components P40m, N100m, and P200m "could be explained by cortical activity within the Sylvian fissure" (p. 31). In this study, Hari and colleagues further demonstrated that the neural sources generating the magnetic counterpart of the P2, the P200m, were located between 9 to 10 mm anterior and 5 mm medial to sources generating the magnetic counterpart of the N1, the N100m.

Waveform Components

Latency and amplitude parameters are important when analyzing waveform components. In normal hearing adult listeners, the first component of the P1-N1-P2 complex is a positive deflection (i.e., P1) occurring between approximately 40 and 150 ms following stimulus onset. The N1 response occurs between 75 and 200 ms and is followed by a positive component, P2, occurring at approximately 175 ms (Wood & Wolpaw, 1982). In terms of amplitude, Kraus,

McGee, Carrell, Sharma, Micco, and Nicol (1993) regarded P1, N1, and P2 as present if these amplitude values are equal to or exceed 0.5 μV.

Central Auditory Processing Evaluation

With respect to P1, N1, P2 waveform components, it has been proposed that these responses may be an effective electrophysiological means of evaluating central auditory processing (Ponton, Vasama, Tremblay, Kwong, & Don, 2001; Tremblay, Kraus, McGee, Ponton, & Otis, 2001). More specifically, P1-N1-P2 components are suggested to reflect stimulus representation in the central auditory system (Martin, Sigal, Krutzberg, & Stapells, 1997; Ostroff, Martin, & Boothroyd, 1998; Sharma & Dorman, 1999, 2000; Whiting, Martin, & Stapells, 1998) and sensory encoding of auditory stimulus characteristics (Näätänen & Picton, 1987). Generally, the P1-N1-P2 complex, specifically the N1 component, is thought to represent that a signal is perceived at the level of the auditory cortex and can be used to measure central auditory processing abilities especially in the temporal domain.

Ponton, Vasama, Tremblay, Kwong, and Don (2001) recorded N1-P2 responses in adults with normal hearing and those with a unilateral deafened ear. They reported that compared to normal hearing listeners those with unilateral hearing loss demonstrated increased P1-N1 peak-to-peak amplitudes when recorded from the ear without hearing loss (i.e., the normal hearing ear). "Significant increases in inter-hemispheric waveform cross-correlation coefficients, and in inter-hemispheric AEP [Auditory evoked potentials] peak amplitude correlations" (p. 32) were also demonstrated in responses recorded from the group with unilateral hearing loss compared to the participant group with bilaterally normal hearing. Ponton et al. (2001) determined that unilateral hearing loss affects the central processing of auditory stimuli and these effects are represented in the N1-P2 recording. In light of these findings Tremblay et al. (2001) suggested

that the N1-P2 is possibly an effective measure of central auditory processing. Again, in adults with normal hearing and normal auditory processing, P1-N1-P2 recordings should be composed of a positive-negative-positive complex occurring between 50 to 250 ms (Stapells, 2002) with peak-to-peak amplitudes between 0.5 to 5 μV.

Stimulus Effects

Intensity. Researchers have established that stimulus parameters such as intensity, frequency, and duration affect N1-P2 responses; however, little work has been done to identify the effect on the P1 component. Therefore, this review mainly focuses on the stimulus effects for the N1-P2 components. For example, in earlier investigations it was reported that there is an inverse relationship between amplitude and latency as a function of stimulus intensity. In that, as stimulus intensity increases, N1-P2 amplitude increases and N1-P2 latency decreases (Antinoro, Skinner, & Jones, 1969; Beagley & Knight, 1967; Davis, Mast, Yoshie, & Zerlin, 1966; Davis & Zerlin, 1966; Gerin, Pernier, & Peronnet, 1972; Onishi & Davis, 1968; Picton, Woods, Barbeau-Braun, & Healy, 1977; Rapin, Schimmel, Tourk, Krasnegor, & Pollark, 1966; Rothman, Davis, & Hay, 1970). This phenomenon was shown to occur until approximately 70 dB, at which point as intensities are increased beyond this level the effect on amplitude and latency was not as pronounced (Geisler & Murphy, 2000; Picton et al., 1977; Rothman et al., 1970; Sugg & Polich, 1995). In other words, at intensities above 70 dB, increases in amplitudes and decreases in latencies taper off. Notably, Kasey, Salzman, Klorman, & Pass (1980) showed the same trend of amplitude increases with intensity increases for P1-N1 complex recorded using tone burst presented at 60, 77, and 95 dB.

Researchers have also looked at N1 and P2 waveform components separately and have found differing response patterns in terms of intensity effects. Adler and Alder (1989) presented

1000 Hz tone bursts to 10 normal hearing adult listeners at 13 intensities from 30 to 90 dB SL in 5 dB steps. In order to separate N1- P2 waveform components Adler and Adler measured N1-P2 amplitude as a peak-to-baseline measurement instead of the traditional trough-to-peak measurement. For intensities between 30 to 70 dB SL reported findings were consistent with other investigators in that both N1 and P2 amplitudes increased in a linear manner. At intensities above 70 dB, it was reported that N1 amplitude began to decrease, where as P2 amplitude gradually increased. Alder and Alder (1989) also demonstrated differences between N1 and P2 latencies as a function of intensity, in that, as intensity increased between 30 and 70 dB SL P2 latency shifts were much larger than N1 latency shifts. Additionally, other investigators have reported an inverse relationship between latency and intensity (i.e., latency decreases as intensity increases) up to approximately 70 dB (Beagley & Knight, 1967). Alder and Alder conversely reported a "U-shaped relationship" between latency and intensity for both N1 and P2 components with those evoked via the 70 dB SL tone producing the minimum latency values.

Duration. Stimulus duration is a highly investigated stimulus characteristic with respect to the effect on P1-N1-P2 waveform components. Some of the ALR seminal studies focused on various aspects of duration (Davis & Zerlin, 1966; Onishi & Davis, 1968). Onishi & Davis (1968) conducted two experiments, which investigated the effect of stimulus rise/fall times, duration, and intensity on N1, P2 components. In the first experiment, 1000 Hz tone bursts were presented at six durations (i.e., 0, 3, 10, 30, 100, and 300 ms) with one of two rise/fall times (i.e., 3 or 30 ms) at four intensities (i.e., 25, 45, 65, 85 dB); totaling 48 variations of the 1000 Hz tonal stimuli. In the second experiment, 18 stimuli were used consisting of the 1000 Hz tone presented at six rise/fall times (i.e., 3, 10, 30, 50, 100, and 300 ms) a constant 2.5 s duration at three intensity levels (i.e., 45, 65, and 85 dB). With respect to experiment one, Onishi and Davis

reported that for the 1000 Hz tone bursts where the rise/fall time was 30 ms, amplitude was not affected as the duration was increased from (0-300ms). For those tones with a 3 ms rise/fall time an increase in N1-P2 amplitude was noted at shorter durations (i.e., 0-30 ms); however, at plateau durations longer than 30 ms amplitude was not affected. Given this, it was suggested that N1-P2 amplitude is determined within the first 30 ms post stimulus onset. N1 latency was also not affected by durational changes in tones with a 30 ms rise/fall time. For 3 ms rise/fall tone bursts, a prolongation of N1 latency was reported at shorter plateau durations with the most notable increase being the prolongation of N1 latency when elicited with a 3 ms rise/fall time and a 3 ms plateau at 45 dB.

Investigators have further focused on the effects of stimulus duration on the temporal integration time represented via auditory ERP latencies. According to Hall (2007), "temporal integration time, as assessed with auditory evoked responses, corresponds to the minimum duration of a signal that produces maximum AER amplitude. Among ERPs, temporal integration times are directly related to latency of the response" (p. 493). Eddins and Peterson (1999) recorded N1 and P2 components using 1000 and 4000 Hz tone bursts at durations of 8, 16, 32, 64 and 128 ms and intensities of 10, 20, 30 and 40 dB SL. In this investigation it was reported that as the duration of the tones increased N1 and P2 latencies decreased. However, similar to findings by Onishi and Davis (1968), these durational effects were most notable when the tone bursts were presented with shorter durations (i.e., 8-32 ms). For the tone bursts with longer durations (i.e., 64 and 128 ms) similar component latencies were observed. Durational effects on N1-P2 amplitude were not reported.

As related to temporal integration, numerous studies have evaluated the effects of stimulus duration as related to intensity. It is well established that stimulus duration affects

loudness. As the duration of the stimulus, decreases the perceived loudness of the stimulus also decreases (Moore, 1997). Equivocal findings have been reported regarding the precise relationship between duration and intensity. A relatively linear relationship has been demonstrated for stimuli with durations less than 200 ms where the effect of temporal integration of loudness is suggested as being the most pronounced (Garner & Miller, 1947; Green, Birshall, & Tanner, 1957). However, at stimulus durations of greater than 200 ms the effect of duration is minimal.

Conversely, Eddins and Peterson (1999) reported that increasing stimulus duration between 64 ms and 124 ms did not produce a significant decrease in N1-P2 latencies as did durational changes between 8 and 64 ms. These authors found N1-P2 latency effects were only significant with respect to duration at low level intensities close to threshold (i.e., 10 and 20 dB SL). At intensities above threshold (i.e., 30 and 40 dB SL) N1-P2 latency shifts were not observed for either the 1000 or 4000 Hz tone.

Agung, Purdy, McMahon, and Newall (2006) reported equivocal findings. They investigated the influence of duration on P1-N1-P2 wave components using seven consonant and vowel tokens at durations of 100 ms and 500 ms presented via loudspeaker at 65 dB SPL. N1 and P2 component amplitudes were significantly larger when recorded using shorter duration tokens (i.e., 100 ms) than those recorded using the longer 500 ms tokens. With respect to latency, significantly earlier latencies were found for P1-N1-P2 components for the 100 ms vowels.

Frequency. Numerous investigators have reported a frequency effect for ALR waveform components. When elicited via a low frequency (i.e., 250-400 Hz) stimulus amplitudes increase and latencies decrease compared to higher frequency stimuli (i.e., 1000-3000 Hz; Alain, Woods & Covarrubias, 1997; Antinoro et al., 1969; Jacobson, Ahmad, Moran, Newman, Wharton, et al.,

1992; Sugg & Polich, 1995; Wunderlich & Cone-Wesson, 2001). Crottaz-Herbette and Ragot (2000) evaluated N1 responses under nine stimulus conditions in eight males and eight females. These stimuli consisted of one of three fundamental frequencies (i.e., 62.5, 125, and 250 Hz) varied with three different spectra (i.e., low, mid, and high). Crottaz-Hebette and Ragot reported a stimulus effect, in that N1 latencies decreased as fundamental frequency increased. Shestakova, Brattico, Soloviev, Klucharev, and Huotilainen (2004) and Agung et al. (2006) reported similar findings.

Agung and colleagues (2006) suggested that the influence of frequency on cortical potentials is possibly the result of the tonotopicity of the auditory cortex. Yetkin, Roland, Christensen, and Purdy (2004) demonstrated that low frequency stimuli activated more superficial cortical areas than high frequency stimuli. The authors reasoned that response amplitudes for low frequency stimuli are larger when recorded from scalp electrodes because this superficial activation pattern is produced and is more easily recorded from the electrode.

Interstimulus interval. Interstimulus interval (ISI) is defined as the "time between successive stimulus presentations" (Stach, 2003, p. 143). Earlier researchers reported that as ISI increased (i.e., stimulus presentation rate slowed) N1 and P2 wave amplitude also increased; however, ISI variations did not affect N1 and P2 latencies (Davis, 1965; Davis et al., 1966; Davis & Zerlin, 1966; Fruhstorfer, Soveri, & Jarvilehto, 1970; Hari et al., 1982; Keidel & Spreng, 1965; Nelson & Lassman, 1968; Picton, et al., 1977; Rothman, Davis, & Hay, 1970). Although, typically related to response latency it was concluded that ISI effects on component amplitude were the result of auditory cortical neural refractory periods (Hall, 2007). Neural refractory periods are related to the time that it takes a neuron to return to a state of rest post stimulation, at which point, the neuron is able to maximally respond to additional stimulation. If the neuron is

stimulated prior to returning to its resting state it will be unable to provide this maximum response. Longer ISI are necessary for the recording of auditory long latency ERPs and typically ISI of at least 1/s are utilized.

Tonal stimuli versus speech stimuli. How does stimulus type (e.g., tonal vs. speech) affect long latency response recordings? Given that the P1-N1-P2 complex provides an electrophysiological method of evaluating neuronal processing and is sensitive to acoustic cues it stands to reason that responses will be different when elicited with single frequency tones versus speech stimuli. To investigate this question, many researchers have examined long latency ERPs to various stimuli including complex tones (Martin & Boothroyd, 1999), natural and synthetic vowels (Diesch, Eulitz, Hampson, & Ross, 1996; Eulitz, Diesch, Pantev, Hampson, & Elbert, 1995; Gunji, Hoshiyama, & Kakigi, 2000), and syllables (Ostroff, Martin, & Boothroyd, 1998).

Several differences in the P1-N1-P2 responses recorded via tonal stimuli and speech stimuli have been reported. Namely, responses recorded via tone bursts have a shorter P1-N1-P2 latencies and smaller amplitudes than those elicited via speech stimuli (Cëponiene, Shestakova, Balan, Alku, Yiaguchi, & Näätänen, 2001; Tiitinen, Sivonen, Alku, Virtanen, & Näätänen, 1999).

Tiitinen and colleagues (1999) employed MEG technology to simultaneously record ERPs (i.e., N1-P2 responses), and the magnetic components (i.e., N1m – P2m), to both tonal and vowel stimuli. In order to match the tonal stimuli and speech stimuli in terms of frequency and intensity, Tiitinen and colleagues used a 535 Hz tone and two Finnish vowels /a/ and /A/. It was reported that N1-P2 response amplitudes were larger for the vowel stimuli than the pure tone stimuli with no amplitude differences between the two vowels (i.e., /a/ and /A/). In terms of N1 latency, Tiitinen et al. reported that latencies were shorter for the tonal stimuli than the /a/ vowel

stimuli, but the tonal stimuli were not significantly different from the /A/ stimuli. There was a significant difference between vowels /A/ and /a/ with N1 latencies being significantly shorter when recorded with vowel /A/ compared to the vowel /a/. Further, it was reported that P2 latencies recorded with the pure tone were shorter than those recorded with both vowel stimuli and there were no reported differences between responses recorded with vowel /a/ compared to vowel /A/. When analyzing N1-P2 latency data as a whole it was concluded that responses elicited using a pure tone differ from those elicited with both vowels.

Cëponiene et al. (2001) recorded P1, N250, and N450 wave components to synthetic vowel tokens, complex tones, and tone bursts from 8-10 year old children to investigate stimulus effects. They found that P100 and N250 amplitudes were larger when recorded with complex tonal stimulus compared to the "sinusoidal tones." The N250 component amplitude was also smaller when recorded using the vowel token relative to the complex tones, but the N450 component amplitude was larger when recorded with the vowel token compared to the complex tones. With respect to latency differences, P100 latencies were statistically shorter for the complex tones compared to the vowel token. No P100 latency differences were found between the vowel and the sinusoidal tones conditions. N250 latencies were shorter for the sinusoidal tones compared to the vowels, but not between the complex tones and the vowel tokens. No significant N450 latency differences were reported between any of the conditions.

Equivocal findings have also been reported with respect to response latencies for tone and speech tokens, in terms of the frequency of the eliciting stimulus. As described above numerous investigators have reported a frequency effect on responses recorded with tonal stimuli. Responses recorded with tonal stimuli tend to vary as a function of the frequency of the eliciting stimuli (Crottaz-Herbette & Ragot, 2000). Mäkelä, Alku, & Tiitinen, (2003) however,

reported that N1m response latencies were not affected by the frequency differences between the first two formants of vowels /a/, /o/, and /u/. It was reported that N1m latencies for all tokens were approximately 120 ms. Contradictory to these findings, Agung and colleagues (2006) reported N1 and P2 latency shifts for naturally produced tokens were similar to those reported for tonal stimuli (i.e., shorter component latencies for tokens with more low frequency energy).

Researchers have shown that long latency ERP wave components are affected by other temporal parameters of the speech stimuli such as voice onset time (VOT) and voicing characteristics (i.e., voiced vs. voiceless; Kutzberg, 1989; Sharma et al., 2000; Steinschneider et al., 1999; Tremblay, Piskosz. & Souza, 2003). VOT is the time between articulatory release and the onset of voicing. With respect to long latency ERPs, researchers have demonstrated larger component amplitudes when recorded via voiced speech sounds compared to those recorded via voiceless speech sounds and these findings held true for both synthetic speech tokens (Steinschneider et al., 1999) and natural speech tokens (Tremblay et al., 2003). In a more recent study, Digeser, Wohlberedt, and Hoppe (2009) also found that N1 and P2 component latencies were shorter and component amplitudes were larger when recorded via a naturally produced /ta/ token compared to responses recorded via a naturally produced /da/ token.

Another notable difference between tone elicited and speech-elicited responses is the demonstrated differences in hemispheric lateralization. Szymanski, Rowley, and Roberts (1999) found that the MEGs recorded using a 1000 Hz tone and synthesized speech /a/ and /u/ tokens differed in terms of hemispheric asymmetry. Hemispheric activity was larger in the left hemisphere than in the right hemisphere when recorded with speech, while, symmetrical hemispheric responses were produced when responses were elicited with tonal stimuli. The authors also reported that M100 component latency did not differ in terms of tone vs. speech

stimuli. However, M100 component amplitudes were larger for the tones than the vowel tokens. Mäkelä et al. (2003) also reported left hemispheric lateralization of vowel tokens.

Natural speech tokens versus synthetic speech tokens. Is there a difference between P1-N1-P2 complex latencies and amplitudes when elicited via naturally produced speech tokens in comparison to synthetic synthesized speech tokens? Picton, Alain, Otten, and Ritter (2000) suggested that naturally produced speech should be used in electrophysiological testing so that the results can be related to everyday listening experiences. However, some authors argue that synthesized speech tokens allow for the manipulation of distinct acoustical cues (Digeser et al., 2009). Although, little work has been done to investigate the differences between neuronal responses elicited using natural versus synthetic speech tokens, auditory long latency evoked potentials have been reliably recorded using natural speech.

Ostroff, Martin, and Boothroyd (1998) recorded the N1-P2 complex using the natural produced syllable [sei], the [s] consonant sound taken from the syllable, and the [ei] vowel portion of the syllable. It was reported that N1-P2 complexes were elicited with all three stimuli; however, the [ei] vowel stimulus produced N1-P2 complexes with larger amplitudes than the consonant [s] stimulus. Further, when subtracting the [s] waveforms from the whole syllable [sei], the resulting waveform did not match that of the isolated vowel [ei] waveform. Based on these results, Ostroff and colleagues concluded that the N1-P2 complex might be reflective of the "underlying acoustic patterns" of the stimuli.

Tremblay, Friesen, Martin, and Wright (2003) recorded P1-N1-P2 complexes via naturally produced speech tokens /bi/, /pi/, /shi/, and /si/. The aim of this investigation was to determine if reliable and replicable P1-N1-P2 complexes could be recorded using naturally produced tokens. Tremblay, Friesen et al. (2003) questioned this, given, the variability and

complexity of naturally produced speech and speculated that naturally produced speech may not be an ideal stimulus for recording the neuronal processing of the acoustic cues of speech such as VOT. However, these authors did suggest that if replicable responses were recorded then the use of natural speech tokens might be more beneficial than synthesized speech for the assessment of speech processing. Seven normal hearing listeners served as research participants. Responses were recorded over two test sessions over a time period of eight-days. Individual responses were not significantly different between test sessions, leading to the conclusion that natural speech was an appropriate stimulus for recording P1-N1-P2 complexes. These authors also reported differences in N1 and P2 amplitudes related to VOT of the tokens (i.e., N1 and P2 amplitudes were larger when recorded with the /bi/ token than the /pi/ token). Notably, the voicing aspects of the tokens (i.e., /bi/ vs. /pi/) did not affect N1 and P2 latencies. In previous studies presenting synthesized tokens, component latencies were affected by voicing aspects, in that, earlier N1 latencies were recorded with voiced tokens than with voiceless tokens (Sharma, Marsh, & Dorman, 2000; Tremblay et al., 2002). According to Tremblay and colleagues, this finding was important and may indicate differences in the central processing of synthesized and natural speech tokens.

Gender Differences

Gender differences have been noted across numerous electrophysiological indices and have been extensively investigated in early-evoked responses, namely, ABRs. ABRs in female listeners compared to male listeners typically demonstrate larger wave V amplitudes (Chan, Woo, Hammond, Yiannikas, & McLeod, 1988; Dehan & Jerger, 1990; Jacobsen, Novotny, & Elliot, 1980; Jerger & Hall, 1980; Michalewski, Thompson, Patterson, Bowman, & Litzelman, 1980; Trune, Mitchell, & Phillips, 1988; Sand, 1991), shorter wave V latencies, and shorter wave

I-V latency intervals (McClelland & McCrea, 1979; Rosenhamer, Lindström, & Lundborg, 1980; Bergholtz, 1981; Elberling & Parbo, 1987; Thornton, 1987; Durrant, Sabo, & Hyre, 1990; Fujita, Hyde, & Alberti, 1991) for ABRs elicited with click stimuli.

Although not extensively researched, equivocal findings have been reported with respect to gender differences when recording auditory long latency response (Hall, 2007; Hyde, 1997). For example, Polich, Howard, and Starr (1985) did not report gender differences when recording P300 event-related potentials with an oddball paradigm in the presence of a masking noise. Onishi and Davis (1968) examined the effects of tone burst duration and rise fall time on auditory ERPs in both male and female participants. They reported that female participants demonstrated overall higher amplitudes and steeper intensity functions than male participants. Altman and Vaitulevich (1990) reported that shorter N1 latencies were recorded from female participants than male participants. Golgeli, Suer, Ozesmi, Dolu, Ascioglu, and Sahin (1999) used an oddball paradigm with a 2000 Hz standard and a 1500 Hz target stimuli to record ERPs (i.e., N1, P2, N2, and P3) from 20 male adults and 18 female adults. In this study, male subjects demonstrated higher N1-P2 and N2-P3 interpeak amplitudes at the Cz electrode site and higher N2-P3 interpeak latencies at the Oz electrode site than females. N1, P2, N2, P3 latency differences were not observed.

Differences in hemispheric lateralization have also been reported between males and females. Rockstroh, Kissler, Mohr, Eulitz, Lommen, Wienbruch, et al. (2001) reported that male participants displayed greater activity in the left-hemisphere when elicited with /ba/ tokens than female participants. Obleser, Eulitz, Lahiri, & Elbert (2001) demonstrated greater left hemispheric activity in female participants than male participants during vowel discrimination. Obleser, Rockstroh, and Eulitz (2004) further investigated hemispheric processing differences

between male and female participants. N1m responses were recorded using naturally produced speech tokens and noise stimuli for comparison. Further, participants were asked to behaviorally identify syllable tokens based on the initial 35 ms of onset. The authors reported that female participants showed increased N1m responses in the left hemisphere than in the right hemisphere and those female participants who displayed greater left hemispheric lateralization also performed better during the categorization task. Male participants did not demonstrate left hemisphere dominance for speech tokens and left hemisphere dominance was not an indicator for syllable categorization performance. Further, hemispheric lateralization was only present for speech-evoked responses. In light, of these findings Obleser and colleagues concluded that speech processing differs between male and females.

Brain Imaging Techniques

Advancements in functional neuroimaging techniques have also allowed for the measurement of the subtle temporal and spatial characteristics of cortical and subcortical structures underlying speech production and language perception (Kertesz, 1994). Although, imaging studies are not the focus of the present investigation, a fundamental understanding is necessary given numerous studies have utilized these recording techniques to examine audition during speech production. Thus, it is essential that a brief review of brain imaging techniques be presented below.

Magnetoencephalograhy

MEG is a noninvasive neuroimaging technique that offers excellent spatial resolution on the superficial fissures (i.e., on the order of 1 mm) and the basal temporal and frontal areas (Tarkiainen, Hämäläinen, & Salmelin, 1997; Leahy, Mosher, Spencer, Huang, & Lewine, 1998) as well as excellent and temporal resolution (i.e., on the order of 1 ms). Therefore, determining

the location and timing of cortical generators is the primary objective of recording MEGs (Paetau, 2002). MEG technology detects and records these magnetic fields emitted during neuronal processing and displays them as a series of traces that can be analyzed. Importantly, MEG is less susceptible to artifacts than EEG and provides a tool for tracking brain activity and disruptions during a variety of tasks (Salmelin, Hari, Lounasmaa, & Sams, 1994; Helenius, Salmelin, Service, & Conolly, 1998).

Electroencephalography

EEGs were one of the first noninvasive techniques developed for observing brain activity. Hans Berger developed this technique in the early 1900s. He discovered that spontaneous electrical energy produced by the human brain can be measured as a function of time and in doing so allows researchers to observe how this electrical energy changes in response to presented stimuli. This electrical energy is presented in the form of brain waves and is plotted on an electroencephalogram.

There are four classifications of brain waves recorded by EEGs. These brain waves are defined in terms of frequency and are labeled as alpha, beta, theta and delta waves. Alpha waves occur between eight and 13 Hz and are elicited in individuals who are relaxed and have their eyes closed. These waves can be used as a measure of hemispheric asymmetry. The presence of high amplitude alpha waves indicates the interruption of organized thoughts (Morgan, McDonald, & MacDonald, 1971). Conversely, when an individual allocates attention to a presented stimulus the alpha activity will be suppressed (Adrian & Matthews, 1934). Therefore, comparing the hemispheric alpha suppression allows for the inference of which hemisphere is processing the presented stimuli (i.e., allows for the determination of hemispheric asymmetry).

EEGs waves can be employed to determine the overall state of the brain and are commonly used to identify abnormal electrical activity in disorders such as epilepsy. For example, beta waves occur between 13-30 Hz in individuals who are alert or have taken medications such as benzodiazepines. However, beta waves will be reduced or even absent if cortical damage is present. Delta waves occur at 3Hz or less. These waves are typically present in infants and young children or in sleeping adults. Delta waves are sensitive to cortical injuries or encephalopathies such as hydrocephalus. Theta waves occur between four to seven Hz and are typically only seen in children or sleeping adults. Therefore, cortical injuries may be present if theta waves are recorded in adults who are awake.

EEGs recorded in individuals without cortical damage will demonstrate the following characteristics: alpha and beta waves primarily recorded in alert individuals; stable electrical activity throughout the test and symmetrical electrical brain activity in the absence of stimuli (i.e., visual or auditory). Indices of cortical damage through EEG recordings include: asymmetric activity in the absence of stimuli; sudden bursts of electrical activity; the presence of delta or theta waves in alert adults; and the absence of electrical activity (i.e., flat brain waves). Although, EEGs recordings allow good resolution in the temporal domain, recordings are limited in terms of spatial resolution. EEGs represent the electrical activity from a large number of neurons as a function of seconds and do not show the specific anatomical location where this activity is generated.

Positron Emission Tomography

Positron emission tomography (PET) is a nuclear imaging technique that produces functional images of neuroanatomical structures and neurophysiological processes. PET scans measure cerebral blood flow and the changes in cerebral blood flow in response to various

stimuli. In order to measure these changes a radioactive substance or tracer is injected into the body. Increases in cortical processing will cause an increase in the amount of regional cerebral blood flow. PET scans detect this increase of regional cerebral blood by showing increases in the amount of radioactive substances in cortical structures.

Indices of Motor-to-Sensory Integration

Again, technological advancements have allowed for the non-invasive investigation of speech production and speech perception. Following is a summary of studies employing electrophysiological measures, such as, ERPs and MEGs, to examine neural responses as related to speech production and speech perception. It should be noted that the majority of this research was performed to examine differences in auditory cortical activity (i.e., cortical suppression) between individuals diagnosed with and without schizophrenia.

Electrophysiological Indices

Kudo, Nakagome, Kasai, Araki, Fukuda et al. (2004) investigated the notion of corollary discharge suppression during speech production through the recording of N1, mismatch negativity (MMN), negative difference wave referred to as Nd, and P300 ERPs in twenty two (i.e., 11 males and 11 females) normal right-handed adults. Initially, a baseline condition was recorded using an "odd ball" paradigm. In this condition, participants were presented with 1000 Hz and 2000 Hz tone bursts that varied in duration between 25 ms and 100 ms and were asked to press a button when hearing the target stimulus in the ear specified as the attended ear. Upon completion of this baseline task (i.e., where both right and left ears served as the attended ear), participants were asked to perform the same listening task while vocalizing the vowel /a/. The results of this study can be summarized as follows: suppression of N1 amplitude was noted during the vocalization task compared to the baseline condition; this N1 amplitude suppression

was not found during any of the other ERPs (i.e., MMN, Nd, or P300); no significant gender effects were reported.

Although, Kudo et al., (2004) concluded that their findings support the notion of auditory corollary discharge, given that N1 amplitude suppression was demonstrated, these authors questioned why similar suppression patterns were not seen during the other ERP measures. With respect to the function of corollary discharge signals, Kudo et al. speculated that corollary discharges might only function at the perceptual level. Further, these authors suggested that dividing participant attention between the tasks of vocalizing and selectively listening to the incoming auditory signals might affect ERPs (e.g., Nd), which are influenced by listener attention. Consequently, the simultaneous recording of multiple ERP components possibly affected MMN responses because the optimal recording methodology was not utilized.

The notion of sensory-to-motor matching and the relationship between corollary discharge signals and inner speech/thoughts has also been extensively researched in individuals with schizophrenia (Bentall, 1990). Namely, researchers have posited that auditory hallucinations are the result of deficits in the corollary discharge mechanisms. That is, there is a disruption in the corollary discharge, which results in the schizophrenic individual perceiving internal thoughts as externally generated. Feinberg (1978) and Frith and Done (1989) supported this notion and hypothesized that deficits in self-monitoring, both auditory and somatosensory self-monitoring, may lead to the behavioral manifestations of schizophrenia (e.g., the generation of auditory hallucinations). These authors suggest that in those with schizophrenia there is a "disconnection" (Friston & Firth, 1995) between the prefrontal lobes and temporal lobes that result in aberrant corollary matching during speaking, which leads to an inability to distinguish self-produced vocalization from external vocalizations. Shergill, Brammer, Amaro, Williams,

Murray, and Phillips (2004) recorded functional magnetic resonance imagines while those with schizophrenia experienced auditory hallucinations. In this study it was shown that the left inferior frontal and right middle temporal gyri were activated between six to nine seconds before the participants noted the onset of an auditory hallucination. During an auditory hallucination activation of the bilateral temporal gyri and the left insula were seen. According to Shergill et al. their findings provide additional support for the "hypothesis that during hallucinations activation in cortical regions mediating the generation of inner speech may precede the engagement of areas implicated in the perception of auditory verbal material" (p. 516) and are consistent with notion of deficient auditory self-monitoring.

Ford, Mathalon, Kalba, Whitfield, Faustman, and Roth (2001) investigated the dampening of N1 event-related potentials (i.e., auditory and visual) during self-produced speech, silent recording, and listening to one's own speech played back in 12 participants with schizophrenia and 10 control participants. In this study, the control group demonstrated a dampening of auditory probe N1 when listening to their speech being played back to them and during the speech aloud condition compared to the silent condition. The 10 participants with schizophrenia did not show a decrease in the auditory probe N1, which may suggest that an absence in N1 reduction is possibly linked to auditory hallucinations. These results were supported by Ford and Mathalon, (2004), Heinks-Maldonado et al. (2005), Heinks-Maldonado et al. (2006), and Heinks-Maldonado et al. (2007) who noted N1 suppression to self-produced speech in normal listeners. Heinks-Maldonado et al. (2005) and Heinks-Maldonado et al. (2007) noted an absence of N1 suppression when elicited with altered vocalizations, such as self-vocalizations that are pitch shifted.

Behroozmand, Karvelis, Liu, and Larson (2009) examined the effect of AAF during passive listening and speech production. P1, N1, and P2 wave components were recorded from 15 right handed participants, while they passively listened to recorded self-produced /a/ tokens, frequency shifted +100, +200, and +500 cents and while the participants overtly vocalized the /a/ token and were presented with FAF (i.e., +100, +200, and +500 cents). The term "cents" refers to divisions or increments of a musical octave. Cents are extremely small increments (i.e., 1200 cents in one octave), which allow for precise fine-tuning. Larger P1 and P2 component amplitudes were reported for responses recording during speaking tasks compared to those recorded during listening tasks and larger P2 amplitudes were recorded using the +100 cents alteration compared to the +500 cents alteration. However, significant N1 component amplitude differences were not seen in this investigation, which is in contrast to previous investigations (Heinks-Maldonado et al., 2005; Heinks-Maldonado et a., 2007). Since increased P1 and P2 amplitudes were found for responses recorded during speech production compared to listening, Behroozmand et al. interpreted results as supporting the notion of a feedback monitoring system and related the absence of an N1 difference to methodological differences between studies.

Brain Imaging Indices

Results from several investigations provide support for a feedforward model of speech production. Numminen, Salmelin, and Hari (1999) recorded MEGs in nine normal speaking participants to investigate auditory cortical activation in response to the participant's own voice. MEGs were recorded under three conditions. The first condition was a control condition in which participants were asked to silently read a provided text while a 1000 Hz pure tone was presented to the left and right ears in an alternating manner. The second condition consisted of

having the participant read a provided text aloud while a 1000 Hz pure tone alternated between the left and right ears. MEGs were recorded under the third condition in only three of the nine participants. This condition consisted of having the participants read aloud while a tape-recording of their own voice was presented binaurally. Numminen and colleagues reported that the M100 response of MEG recordings were delayed 10 to 21 ms and decreased in amplitude by 44 % to 71% during the reading aloud conditions compared to the control condition. They suggested that the M100 suppression is the result of an interaction between speech production and speech perception occurring at a central level.

Curio, Neuloh, Numminen, Jousmaki, and Hari (2000) also investigated how one's own speech affects auditory cortical activation. They recorded MEGs on eight right-handed, fluently speaking participants under "SPEECH" and "REPLAY" conditions. The SPEECH condition consisted of having the participant listen to two tones which differed in frequency (i.e., a standard low tone and a rare high tone) and verbally produce the vowel /a/ in accordance with hearing the standard tone and the vowel /ae/ in accordance with hearing the rare tone. During the REPLAY condition the participant listened to a tape-recording of their responses during the SPEECH condition along with the oddball stimuli. They were asked to mentally compare the two auditory signals and determine if they had performed the initial task correctly. Curio et al. reported three major findings: (a) while speaking, the auditory cortex is activated at approximately 100 ms post-voice onset; (b) in comparison to the REPLAY, a dampening of the M100 response occurs across both the left and right hemispheres during the SPEECH condition; and (c) during self-generated speech, the M100 is more suppressed over the left hemisphere. Based on these results, Cuiro et al. suggested that self-generated speech primes the auditory

cortex and delays auditory cortical reactions. They further suggested that this priming of the auditory cortical areas is related to the self-monitoring of speech.

A Model of Speech Production

Levelt (1983, 1989) dissects the production of spoken words into several stages beginning with the conceptualization of the intended message. During this stage the speaker cognitively formulates or conceptualizes the desired message (i.e., what the speaker wishes to say). Indefrey et al., (2001) referred to this stage as "conceptual preparation." Following, the conceptualization stage a process of "syntactic encoding" ensues, during which, the speaker assigns the appropriate "lemmas" or words to the preverbal and then arranges those words into the correct grammatical order based on previously learned and stored rules. Next, phonological encoding takes place. During phonological encoding the sounds that will ultimately be produced by the articulators are selected.

Although, Levelt (1983, 1989) nicely simplified the steps of language/speech production, the neuroanatomical and neurophysiological structures mediating speech production remain unclearly defined. Historically, Broca's area and more specifically Brodmann Area (BA) 44 has been shown to play an important role in the process of speech motor planning and the sequencing of vocalizations (Nolte, 2002). This notion has been solidified through the study of individuals identified with cerebral damage within Broca's area and behaviorally present with impaired speech production while speech comprehension remained essentially intact (Dronkers, Shapiro, Redfern, & Knight, 1992).

Documented cases of individuals with speech production deficits who do not have lesions within Broca's area have lead to questioning the role of Broca's area in the process of speech motor planning. Dronkers (1996) presented evidence from 25 such individuals. For these 25

participants, speech-planning deficits were behaviorally manifested, however, the common site of lesion between participants was not Broca's area, but lesions in the left precentral gyrus of the insula. Dronkers suggested that the left precentral gyrus of the insula may be involved in planning and execution of speech; while Broca's area may contribute more to the process of syntactic encoding (Dronkers, Redfern, & Knight, 2000).

In normal participants, insular involvement in speech production has been documented using various imaging techniques including position emission tomography scans (e.g., Wise, Greene, Buchel, & Scott, 1999), functional magnetic resonance imaging studies (e.g., Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000), and MEG studies (e.g., Kuriki, Mori, & Hirata, 1999). Notably, Riecker et al., (2000) reported that while participants silently spoke or sang, only seen regions of the motor cortex and the cerebellum were activated; whereas, the left anterior insula did not become active until participants overtly vocalized.

In addition, researchers have pointed to regions of the premotor cortex, namely the supplementary motor area (SMA; BA 6) as involved in the sequencing and planning of speech production. Chung, Han, Jeong, and Jack (2005) performed magnetic resonance imaging (MRI) on 16 right-handed participants as they performed tasks designed to facilitate certain aspects of speech production (i.e., word generation), language comprehension, and memory. Also, to examine the SMA activation during non-speech related tasks; MRIs were recorded during a finger tapping condition and a heat sensory condition. In terms of SMA activation, the results were reported as follows: (a) during the finger tapping condition SMA activation was seen in all 16 participants, (b) in 15 participants during the heat sensory condition, (c) in 15 participants during the word generation condition, (d) in five participants during the listening comprehension condition, and (e) in 15 participants during the working memory condition. These authors also

noted that, although not significantly, the rostral and caudal portions of the SMA responded differently during the various conditions. More activation from rostral portions of the SMA was seen while the participants were engaged in word generation and working memory conditions, whereas, the finger tapping and heat sensory condition lead to more caudal activations.

These SMA findings reported by Chung and colleagues (2005) were consistent with those reported by previous investigators, suggesting that regions of the SMA function in a task dependant manner. This notion has been previously investigated in both human (Alario, Chainay, Lehericy, & Cohen, 2006; Vorobiev, Govoni, Rizzolatti, Matelli, & Luppino, 1998) and primate models (Bates, & Goldman-Rakic, 1993; Luppino & Rizzolatti, 2000) and a substantial pool of data has emerged that the SMA is composed of multiple regions and each region is specialized for certain tasks. In humans, it is typically accepted that the SMA is composed of two parts, the pre-SMA and the SMA proper (Alario et al., 2006). It has been suggested that the pre-SMA is predominately involved with executive functions such as the motor planning and sensory integration of sequenced motor tasks (Ferrandez, Hugueville, Lehéricy, Poline, Marsault, et al., 2003; Halsband, Ito, Tanji, & Freund, 1993; Hikosaka, Sakai, Miyauchi, Takino, Sasaki, et al., 1996) and accordingly has projections to the prefrontal cortex (Bates & Goldman-Rakic, 1993), which is involved in functions such as working memory and decision making (Notle, 2002). The SMA proper; however, is considered to be more involved in the execution of motor tasks, given that projections from this area go the primary motor cortex and spinal cord (He, Dum, & Strick, 1995).

For the purposes of the present investigation, experimental paradigms were implemented in the context of Guenther and colleagues' previously described DIVA model. A similar model, albeit with less anatomical detail, has also been described by Heinks-Maldonado et al., 2005,

2007 and will also be referred too below. In a board sense speech production is composed of the conceptualization of an intended message, articulation of the message, and monitoring of the message.

With respect to the schematic representation of the experimental model presented in Figure 1, the phases of speech production (i.e., conceptualization to articulation) are presented, as general concepts in order to simplify the model and not to dismiss the importance of these processes or detailed neuroantomical locations. Accordingly, message formulation, syntactic encoding, and phonological encoding are not independently represented, but are encompassed under the Initiation region of this model (see Figure 1). Simultaneously, speech maps are initiated. Speech maps, described in the DIVA model as "Speech Sound Maps" are the groups of neurons that govern the production and perception of the "speech sounds". Anatomically, functions of the Speech Map have been proposed to take place within Broca's area and the left frontal operculum (Guenther, 2007).

As depicted in Figure 1, once the intended message is constructed and the appropriate sounds and words selected, feedforward commands are projected to articulation maps located in the ventral primary motor cortex (Guenther & Vladusich, 2009). Guenther and colleagues hypothesized that projections from premotor cortical regions to the primary motor cortex and cerebellum mediate the "feedforward control subsystem" and are referred to as "feedforward commands" (Guenther, 2001, 2006, 2007; Guenther & Ghosh, 2003) As defined by Tourville et al. (2008), feedforward control regulates actions based on preexisting models before the motor action is executed. Schematically in Figure 1, solid black arrows represent the synaptic projection of feedforward motor commands. Guenther (2006, 2007) also proposed that the

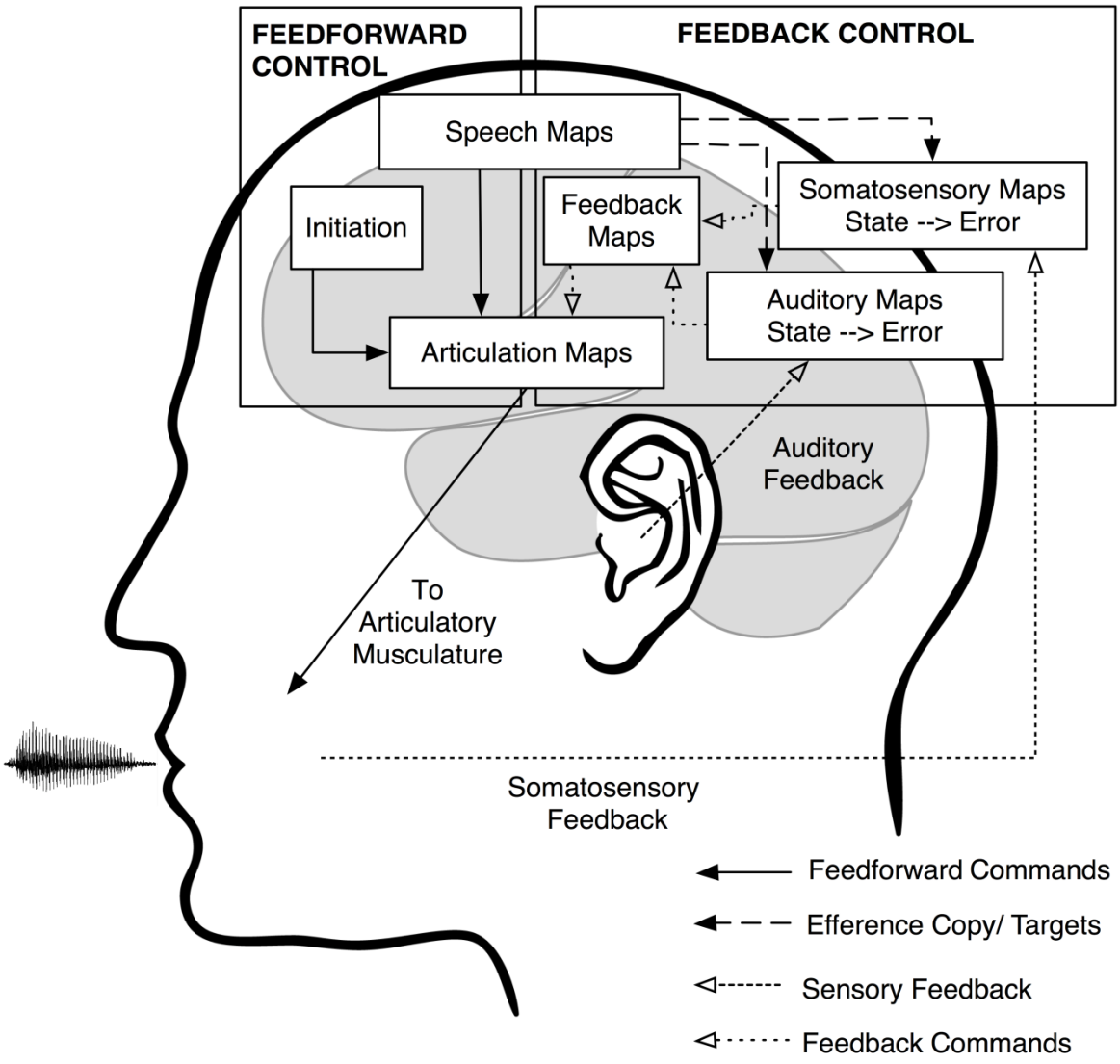


Figure 1. Schematic representation of feedforward and feedback control of speech production. Adapted from "A Neural Theory of Speech Acquisition and Production," by F. Guenther and T Vladusich, 2009, *Journal of Neurolinguistics*, p. 3. Copyright 2009 Elsevier Ltd.

subcortical structure, the cerebellum, aids in the regulation of the rapid timing aspects needed for speech production. The cerebellum is not pictured in Figure 1. At the level of the motor cortex (i.e., articulation maps) the conceptualized message is encoded into the appropriate motor plan that will be executed at the level of the articulators.

At the level of the speech map, efference copies or sensory representations of the motor speech act are simultaneously generated when the speech act is initiated. As depicted in Figure 1, efference copies or sensory targets are projected to somatosensory state and error maps, which according to the DIVA model are located within the supramargnial gyrus (Guenther, 2007) and auditory state and error maps, hypothesized to have contributions from multiple sources such as Heschl's gyrus, superior temporal gyrus, superior temporal sulcus, and the planum temporale (Guenther & Vladusich, 2009). At the level of these maps, the expected feedback is compared with the actual feedback. Heinks-Maldonado et al. (2005, 2007) referred to this expected feedback as corollary discharge. In other words, efference copies are the equivalent sensory copies of the generated motor command. These efference copies are used to generate a sensory representation (i.e., corollary discharge) of the preceding motor actions based on internal and environmental models constructed from previous experiences and stored within cortical working memory areas. In the DIVA model, both auditory and somatosensory feedback control systems are represented. Figure 1 depicts two efference copies are simultaneously generated and projected to the respective map.

When the speech command is executed both auditory and somatosensory feedback or reafference signals (Heinks-Maldonado et al., 2005, 2007) are relayed for further cortical processing (shown in Figure 1). With respect to the auditory feedback control subsystem, overt vocalizations are heard at the level of the peripheral auditory system and then processed through

the central auditory pathway to supratemporal gyrus and planum temporale regions. It should be noted that the supratemporal gyrus region contains cortical areas that are fundamental for auditory processing and speech perception (e.g., the primary auditory [BA 41 and 42] and the auditory association cortex [BA 22] including Wernicke's area; Nolte, 2002). For the somatosensory feedback control subsystem, somatosensory feedback produced by the articulators is projected subcortically to the supramarginal gyrus (BA 40) (Guenther, 2001).

At this point, a comparison is made between the actual auditory input and the corollary discharge (i.e., expected auditory input). Heinks-Maldonado et al. (2005) speculated that "a subtractive comparison of this corollary discharge with the actual sensory feedback associated with the action ("reafference") provides a mechanism for filtering sensory information" (p.180), thus, allowing for a "net cancellation" of cortical activity if there is not a discrepancy between the two signals. In other words, if the corollary discharge and incoming sensory feedback match (i.e., if the two signals cancel each other), a resulting cortical suppression will ensue, notify the speaker that the processed vocalizations are self-generated as opposed to externally generated. In addition, this corollary discharge matching mediates auditory self-monitoring as a way of signaling to the speaker that the intended message was appropriately produced (i.e., in the absence of speech errors) and forward following speech may continue. If the corollary discharges signals are not cancelled by the actual sensory feedback, excitatory input from the incoming stimulus (Guenther, 2007) will initiate corrective actions (Heinks-Maldonado et al., 2005; Guenther, 2006; Tourville et al., 2008). These corrective actions are posited to be produced as a result of a corollary discharge mismatch and are the result of synaptic projections from the auditory error maps, again posited have contributions from multiple sources, including Heschl's gyrus, superior temporal gyrus, superior temporal sulcus, and the planum temporale and the

somatosensory error maps, located with the supramarginal gyrus to the feedback map. The feedback map or feedback control map, located within the right premotor cortex (Guenther & Vladusich, 2009), is hypothesized to further contribute to the encoding of "corrective motor commands" and the selecting of the appropriate feedback commands. The feedback map then projects these selected feedback commands to the articulation maps, which in turn, will initiate the appropriate corrective motor outputs that will be carried out at the level of the articulators. In Figure 1, feedback commands are shown to project from the somatosensory and auditory maps to the feedback map.

Finally, a few additional statements must be made with regard to the current experimental model. Although, not pictured or discussed, significant contributions from parallel processing of the limbic system are made during communication both in terms of internally monitoring self-produced vocalizations and the monitoring of external signals with respect to the emotional aspects of language. Given the complexity of these contributions a detailed account of the limbic system in terms of speech production/perception was beyond the scope of this work. For a discussion see Achermann & Riecker (2004), Gloor (1997), Lamandella (1977), and Larson (1988).

Summary and Research Questions

As presented throughout this review, communication is an intricate process that, for most, seems like an effortless act. The underlying processes of speech production and speech perception; however, are anything but simple. Speech production consists of the rapid formulation of a message, the encoding of that message through the stages of syntactic and phonological encoding, and ultimately the overt articulation of speech sounds (Levelt, 1983, 1989). Amazingly speech is produced at a rate of 4 to 5 syllables/s (Netsell, 1981; Picket, 1980;

Walker & Black, 1950) and involves the integration of the frontal, temporal, and parietal lobes (Greenlee et al., 2004; Guenther, 2006, 2007). For normal speakers/ listeners, the process of communication depends on the ability to transition between being a speaker and being a listener, thus, intricately involving peripheral and central mechanisms regulating both speech production and auditory function (Eliades & Wang, 2008). Even in the most basic exchange, communication partners are required to contribute vocalized ideas or answers, while simultaneously listening to not only to their own voice, but also to the vocalizations of the other individual and to sounds within their surrounding environment.

When considering how speech production and speech perception are used in concert, the interconnection between the two seem apparent. In reviewing the literature what becomes apparent is an obvious discourse surrounding theoretical accounts of speech production and the involvement, if any, of the auditory system in the process. The role of auditory feedback for self-monitoring has historically been a point of debate among researchers, with those opposing auditory feedback involvement subscribing to the notion that auditory feedback is too slow to contribute to rapidly produced speech (Borden, 1979; Levelt, 1983, 1989). Levelt (1989) and Levelt et al. (1999) suggested that an internal perceptual loop analyzes the preverbal message once a phonological code is generated, and this loop is able to detect and correct speech errors 100 to 150 ms before the error is articulated. Further, Levelt and colleagues posited that internally and externally generated vocalizations are cortically monitored via a shared speech comprehension system (Christoffels et al., 2007; Levelt, 1989, Levelt et al., 1999). In other words, self-produced speech is cortically generated, subsequently analyzed for possible errors, vocalized, heard at the level of the auditory periphery, and centrally processed as any incoming auditory input from an external source (i.e., another speaker).

Recently, acceptance for the notion that speech production and speech perception are linked and that auditory feedback is crucial for the monitoring of self-vocalization has substantially grown and a renewed interest in investigating these relationships has emerged. Behaviorally, the influence of auditory feedback is evident when one increases vocal intensity in the presence of competing auditory noise, referred to as the Lombard Effect (Lombard, 1911) and the contrasting lowering of one's voice when competing auditory noise decreases (i.e., the Fletcher Effect; Fletcher, Raff, & Parmley, 1918). Technological advancements such as sophisticated imaging studies have provided neurological evidence supporting sensory-to-motor relationships (e.g., between speech production and speech perception; Ford, Mathalon, Kalaba et al., 2001; Ford & Mathalon, 2004; Guenther, 2006; 2007; Guenther et al., 2006; Hirano et al., 1997; Heinks-Maldonado et al., 2005, 2006; Tourville et al., 2008). However, even with technological advancements, the exact relationship between speech production and perception, including, the exact role of auditory feedback during self-monitoring, has yet to be solidified and theoretical speculation continues.

Given this, plus the lack of unequivocal data supporting currently proposed theoretical models, additional investigations incorporating production–perception tasks seem warranted. The following experiments were conducted for this purpose. Overall, the aim of the current investigation was to further investigate auditory monitoring during speech production and speech perception. These experiments were also constructed to replicate the previous work by Heinks-Maldonado and colleagues and to examine the theoretical assumption of feedforward and feedback control subsystems mediating self-monitoring during speech production.

In order to examine auditory monitoring during passive listening and speech production, auditory ERPs were recorded under various listening and speaking tasks. The incorporation of

passive listening paradigms and speech production paradigms was necessary for two reasons: both are essential for spoken communication and self-monitoring is not only involved in the detection of speech errors, but also allows for the distinguishing of internal and external stimuli. Therefore, examining auditory responses to external stimuli during listening and speaking is essential for a complete investigation of auditory monitoring.

Accordingly, the aim of Experiment 1 was twofold: first, this experiment was designed to address the following general research question: Are auditory ERPs affected by stimulus condition during passive listening? More specifically, are P1-N1-P2 wave components (i.e., latencies and amplitudes) different when recorded under passive listening conditions using non-speech or speech tokens. The listening conditions included, a non-speech tone burst, male and female natural speech tokens, synthetic speech tokens, a recorded participant produced /a/ token, and the recorded /a/ token frequency shifted. Comparisons were made between the non-speech and speech tokens, natural speech and synthetic speech tokens, male speaker and female speaker tokens, the self-produced and other speech tokens, and the self-produced and frequency shifted self-produced tokens. The following hypotheses were tested: First, it was hypothesized that waveform components elicited via the tonal stimulus will differ from those elicited via speech stimuli, due to spectral and temporal differences between the tokens. Second, it was hypothesized that there would be no differences in wave components recorded across speech tokens. An additional aim of Experiment 1 was to provide responses for the comparison with those recorded during active speaking conditions.

Experiment 2 was designed to address these two general research questions: Are auditory ERPs affected by feedback condition during active speaking and Is there a difference between auditory ERPs recorded during speech production compared to those recorded during active

speaking? If the proposed theoretical model holds true, wave components recorded under the NAF condition will have smaller amplitudes than those recorded under AAF conditions. ERPs recorded during overt speaking tasks will also be suppressed or reduced compared to those recorded during passive listening tasks as evidenced by smaller component peak-to-peak amplitudes. In reviewing the literature, it seemed that the predominant method of altering feedback was through the introduction of frequency shifts, with only few investigations utilizing other methods such as DAF (Hirano et al., 1997; Takaso, Eisner, Wise, & Scott, 2010).

In Experiment 2, a FAF condition, a short DAF condition (i.e., DAF 50 ms) and a long DAF (i.e., 200 ms) condition were utilized to investigate the effect of spectral and temporal feedback perturbations on N1-P2 component amplitudes and latency. Comparisons were made for N1-P2 latency and N1-P2 amplitude across feedback conditions. Specifically, comparisons were made between the NAF and FAF conditions, the NAF and DAF 50 ms conditions, the NAF and the DAF 200 ms conditions, the FAF and DAF 50 ms, the FAF and DAF 200 ms, and the DAF 50 and DAF 200 ms conditions. With respect to the four feedback conditions utilized during Experiment 2 the following general hypotheses were tested: It was hypothesized that a feedback effect will be seen across active speaking conditions and a reduction in component amplitude will be seen for the NAF condition compared to the AAF conditions. Specifically, it was hypothesized that no significant differences will be seen for NAF and DAF 50 ms component amplitudes, significant differences will be seen for NAF and DAF 50 ms component amplitudes compared to FAF and DAF 200 ms component amplitudes, and no significant differences will be found for FAF and DAF 200 ms component amplitudes. With respect to ERPs recorded during passive listening compared to those recorded during active speaking it was hypothesized that significantly smaller N1-P2 amplitudes will be recorded during active

speaking compared to passive listening. Heinks-Maldonado et al. (2005) reported that component latencies were not significantly different between passive listening and speaking conditions; therefore, it was hypothesized that component latencies will not be significantly different between passive and active responses recorded here.

Gender differences were also investigated during both experiments. Gender effects for auditory long latency ERPs have not been extensively researched and equivocal findings have been reported across the few studies, which did in fact evaluate P1-N1-P2 components for gender differences. Polich et al., (1985) did not find gender differences for P300 components, were as, Onishi & Davis (1968) reported larger amplitudes and Altman & Vaitulevich (1990) reported shorter N1 latencies for female participants compared to male participants. In light of these findings and given that gender differences are seen across numerous indices of auditory function (e.g., early-evoked potentials; Beagley et al., 1978; Berghotlz, 1981; Chan et al., 1988; Don, Ponton, Eggermont, & Masuda, 1993), it was hypothesized that ERP component latencies will be shorter and amplitudes larger in female participants compared to male participants.

CHAPTER II: AUDITORY EVENT-RELATED POTENTIALS DURING PASSIVE LISTENING

In the words of Ralph Waldo Emerson: “Speech is power: speech is to persuade, to convert, to compel. It is to bring another out of his bad sense into your good sense” (2010, p. 50). In this statement Emerson, rather eloquently, encompasses the enormity of the spoken word. Words are used to express needs and wants, convey emotions, present thoughts and ideas, ask a question or comfort a friend. One’s words have the power to motivate and inspire nations or deliver devastating wounds and for many, hundreds of dollars are spent and countless hours devoted to attending classes and taking seminars to learn how to become more effective communicators.

Essentially, effective communication depends on the ability to rapidly transition between active speaker and passive listener which involves regulating both peripheral and central mechanisms that control speech production and auditory functions (Eliades & Wang, 2008) efficiently and effectively. When considering that both speech production and speech perception necessitate the communication process the connection between these acts seem undeniable. However, the exact role of the auditory system during speech production has been historically questioned. Borden (1979) and Levelt (1983, 1989) discounted auditory system involvement during speech production suggesting that the auditory system is too slow for online speech monitoring and therefore could not contribute to the speech production process. Kalinowski et al. (1993) disproved this notion by showing that auditory input is rapidly processed; therefore, the auditory system is capable of online monitoring of auditory feedback. Recently, acceptance for a speech production and speech perception link has grown and advancements in imaging techniques have provided more sophisticated means of investigating structural and functional

anatomy. Yet, the underlying neurophysiological processes mediating speech production and speech perception remain unclear and speculation continues with the lack of unequivocal data supporting proposed theoretical models. The following experiment was constructed to further investigate neural responses to both speech and non-speech stimulus. In addition, investigations incorporating both production and perception tasks seem warranted due to the intimate relationship between speech production and speech perception; therefore, responses obtained here were obtained during passive listening tasks and will be compared to those obtained during speech production.

Overall, two general research questions were addressed throughout this research. What is the role of audition in the process of speech production and speech perception? What is the link between the motor act of speaking and sensory act of listening? More specifically, this experiment focused on addressing this question: Is there an effect of stimulus and gender on P1-N1-P2 latencies and amplitudes during passive listening?

To investigate neural processing during passive listening P1-N1-P2 components were elicited via the following non-speech and speech stimuli: tone burst, naturally produced male and female /a/ tokens, synthetic male and female /a/ tokens, an /a/ token produced by the participant, and that /a/ token frequency shifted. P1-N1-P2 components were ideal for this experiment given that these tokens are sensitive to changes in the auditory stimulus including competing auditory signals and auditory interference (Hillyard, Hink, Schwent, & Picton, 1973). Sharma and Dorman (1999; 2000) demonstrated that the neural correlates of speech processing could be reliably investigated using auditory long latency ERPs. It was hypothesized that P1-N1-P2 components would differ when elicited using tonal stimulus to compared speech stimulus.

Conversely, it was hypothesized that there would not be significant latency and amplitude differences for all responses recorded with the various speech stimuli.

Gender differences were also evaluated during this experiment. With respect to participant gender the following question was addressed: Is there an effect of participant gender on P1-N1-P2 latencies and amplitudes during passive listening? It was hypothesized that P1-N1-P2 components would differ between male and female participants.

Methods

Participants

This research was approved by the East Carolina University Institutional Review Board prior to data collection or participant recruitment (see Appendix A). Fifteen young adult males and 15 adult females served as participants. Participants were between the ages of 18 and 30 years with a mean age of 24.1 (*SD* = 3.5) years for male participants and a mean age of 24.1 (*SD* = 3.4) years for female participants. All were Caucasian, English speaking, and right handed. Participants were recruited from the East Carolina University student body, the School of Allied Health Sciences student body; and/or a research participant pool located in the Department of Communication Sciences and Disorders. Participants were recruited on a volunteer basis and provided informed consent before testing began (see Appendix B). During the recruiting and data collection process, funding through the East Carolina University Department of Communication Sciences and Disorders became available. The allotted funds totaled $250.00. Five participants were paid a stipend of $50.00 for participating in this investigation. Approval was obtained through the Institutional Review Board to implement this change. The five participants that received compensation signed a separate informed consent before testing began (see Appendix C).

All participants presented with normal-hearing sensitivity defined as pure tone thresholds at octave frequencies from 250 to 8000 Hz ≤ 25 dB HL (American National Standards Institute, 2004) and normal middle ear function defined as Y_{tm} = 0.3-1.50 mmho, TW= 35.80- 95.00 daPa, and V_{ea} = 0.9-1.80 cm^3 (Roup, Wiley, Safady, & Stoppenbach, 1998). Mean hearing thresholds for all participants are displayed in Table 1.

In hopes of decreasing the vast methodological differences seen across electrophysiological investigations, Picton et al. (2000) sought to provide a set of guidelines for the recording and subsequent reporting of data collected using event-related potentials. Picton et al., (2000) also provided suggestions for participant selection and matching. According to these authors, when comparing between group responses, it is essential to establish and document that participants are appropriately selected and matched with respect to participant gender, age, socioeconomic status, and intelligence.

In accordance with these suggestions all participants were first asked to complete a 12 question in take questionnaire requesting age, gender, education level, and total family income; followed by questions concerning the history of neurological disorders, speech and language disorders or learning disabilities. Given that alcohol consumption and certain medications may affect electrophysiological recordings (Picton et al., 2000) this intake questionnaire asked the participant to report if he/she had consumed alcohol or had medications within the 24 hours prior to participation. The complete intake questionnaire is included in Appendix D. Based on the answers provided, it was determined that all participants within this sample were appropriately grouped. All participants were demographically similar with no reported history of neurological disorders, head injuries, otological disorders, learning disabilities, or speech and language impairments.

Table 1.

Participant Mean Hearing Thresholds (db HL) as a Function of Frequency, Ear, and Gender.

	Frequency (Hz)					
	250	500	1000	2000	4000	8000
Male (N = 15)						
Ear						
Right	7.7	6.7	5.0	4.0	6.7	11.7
	(7.3)	(4.9)	(5.0)	(4.3)	(6.5)	(7.2)
Left	7.0	8.3	5.0	6.3	8.7	10.7
	(6.5)	(7.0)	(5.0)	(6.4)	(8.1)	(8.8)
Female (N = 15)						
Ear						
Right	6.3	4.7	5.0	2.7	3.7	12.3
	(4.4)	(5.2)	(4.2)	(6.0)	(6.7)	(7.5)
Left	5.7	5.3	2.7	3.3	1.7	8.0
	(5.0)	(5.5)	(3.2)	(4.1)	(5.8)	(9.3)

Note: Values enclosed in parenthesis represent one standard deviation of the mean.

To assess cognitive function, the Mini-Mental State (MMS; Folstein, Folstein, & McHugh, 1975) was administered to all participants. The MMS is a screening tool of cognitive ability consisting of eleven questions with a total of 30 possible points. Participants were asked each question aloud and scores were calculated as the total acquired points. MMS scores have shown to be affected (i.e., are decreased) by disorders such as dementia, depression, and cognitive impairment. A mean score of 27.6 was reported for normal adults with a range of 24 to 30. Thus, for the purposes of this investigation all participants were required to score 24 points or above on the MMS. All participants demonstrated scores within normal limits (M = 29.90; SD = 0.40). A complete scoring sheet is included in Appendix E.

The Edinburgh Handedness Inventory (Oldfield, 1971) was administered to all participants to assess handedness. The posited relationship between handedness and language make it essential to determine if participants are right or left hand dominant. Typically, motor control (e.g., handedness) and language production/processing share hemispheric laterality and for most, the preferred writing hand represents the dominant hemisphere for motor control. In 90% of individuals, motor functions are controlled by the left hemisphere, which is manifested as right hand dominance (Coren & Porac, 1977) and in the majority of individuals language functions are also dominated by the left hemisphere (Turlough, Fitzgerald, Gruener, & Mtui, 2007). Alexander and Polich (1997) demonstrated a handedness effect on N1, P2, P3 and N2 components. ERPs were recorded from 20 left handed males and 20 right-handed males. It was reported that N1 latencies were shorter and P2 amplitudes smaller in left-handed subjects compared to right-handed subjects. A handedness effect was not demonstrated for N1 amplitudes and P2 latencies. Handedness did not affect the N2 component amplitudes, however, for anterior

electrode sites N2 latencies were shorter in left-handed participants versus right-handed participants.

The Edinburgh Handedness Inventory is a screening measure that is used to examine which hand one uses when performing various tasks. This measure proposes a set of 10 scenarios. The participant was asked to report which hand he/she uses for each of the 10 scenarios. More significance or more points (e.g., double +) are given to either the right or left hand for those tasks where only the preferred hand would be used and the participant would not consider using the non-preferred hand. If the task could be performed using either hand a point or mark (e.g., +) is given to both hands. A laterality quotient (L.Q.) is calculated based on the points given to each hand preference. L.Q.s less than – 40 are considered to represent left-handedness. Score between – 40 and + 40 are representative of those who are ambidextrous. L.Q. scores of + 40 or greater are considered to represent right-handedness. All participants included in the study demonstrated right-hand dominance with L.Q. scores + 40 or greater (M = 72.13; SD = 20.75). A complete scoring sheet is included in Appendix F.

Apparatus

For all experimental conditions, participants were seated in a comfortable reclining chair placed inside a double walled sound treated audiometric suite (Industrial Acoustics Corporation). This audiometric suite met specifications for permissible ambient noise (American National Standards Institute, 1999). Data acquisition was performed using Compumedics NeuroScan 4.4 software, SynAmps^2 Model 8050 EEG amplifier, and Stim^2 software. A 64-Channel Ag/Ag/Cl Compumedics Quik-Cap was used to record ongoing EEG activity. All stimuli were presented via NeuroScan ER-3A insert earphones.

Recording Stimulus

Male and female natural vowel tokens. Natural male and female vowel tokens were recorded for presentation during this experiment. Prior to collecting these vowels, it was first necessary to determine the approximate duration of a typically spoken vowel /a/. To do so, samples of naturally produced /a/ tokens were recorded from ten Caucasian adult male participants with a general American dialect. These speakers were recruited from various departments with the ECU College of Allied Health. All participants were between the ages of 23 and 51 years (M = 29.8; SD = 9.0). All volunteers were asked to generate 10 tokens of the vowel /a/ at a normal vocal effort and normal speech rate into a Logitech (Model 980186-0403) desktop microphone at a sampling rate of 44,100 Hz. Durations of these 100 tokens were determined. Duration was defined as the total length of the vowel, beginning at the point of initial phonation and ending when phonation ceased. On average the duration for these 100 /a/ tokens was approximately 400 ms. A single /a/ token with a duration of 400 ms was selected from the sample of 100 tokens. This token was normalized using Peak 4 software, saved as a .WAV file, and imported into Cool Edit Version 96. In the Peak software the "Normalize" function allows one to "optimize the volume of the selection or an entire document so that it is at its maximum possible amplitude without clipping" (Peak User Guide, p. 141). Additionally, the "Normalize" function allows one to ensure that the amplitude of multiple documents is "uniform". At this point, the vowel was resaved into a format compatible with the Stim2 software program.

The natural female vowel token was recorded from a Caucasian adult female, with a general American dialect. As described above, she was asked to produce the vowel /a/ ten times into the previously described Logitech microphone. This volunteer was also asked to produce the

token at a normal speech rate and a normal vocal effort. The duration of these 10 tokens was determined. Duration was defined as the total time (i.e., ms) of phonation for one vowel token, beginning at the point of initial phonation and ending when phonation ceased. The /a/ token with an approximate duration of 400 ms was selected. The selected vowel was then normalized using Peak 4 software and saved as a .WAV file. This .WAV file was then imported into Cool Edit Version 96 and edited so that the duration of the vowel equaled 400 ms. Once the token was appropriately edited the file was saved as a PC .WAV file that was compatible with the Stim2 AudioCPT module, which was used to present all passive stimuli to participants via insert earphones.

Male and female synthetic vowel tokens. Synthetic speech tokens were created using an Internet based Klatt Synthesizer Interface (http://www.asel.udel.edu/ speech/tutorials/synthesis/) at a sampling rate of 10,000 Hz. The male /a/ was created using the following formant (F) frequencies and bandwidths (BW, all in Hz): F0= 132; F1=710; F2=1150; F3=2700; F4=3300; F5=3700; F6=4900; BW1=40; BW2=43; BW3=105; BW4=400; BW5=500; BW6=1000. These stimulus parameters were chosen based on parameters used by Hawks and Miller (1995). Bandwidths 4 to 6 were based on default settings within the Klatt Synthesizer program. The F0 was varied across the duration of the vowel from 100 to 132 Hz to increase naturalness of the vowel token (Assmann & Katz, 2000). In order to remain consistent with the natural speech tokens, duration of this synthetic token was 400 ms.

The synthetic female /a/ was created using the same Internet based Klatt Synthesizer Interface (http://www.asel.udel.edu/ speech/tutorials/synthesis/) at a sampling rate of 10,000 Hz with the following formats (F) and bandwidths (BW, all in Hz): F0= 208; F1=688; F2=1273; F3=2966; F4=3300; F5=5400; F6=6000; BW1=90; BW2=110; BW3=170; BW4=250;

BW5=300; BW6=450. These stimulus parameters were chosen based on parameters used by Assmann and Katz (2000) and Hawks and Miller (1995). Again, bandwidths 4 to 6 were based on default settings within the Klatt Synthesizer program. The F0 was varied across the duration of the vowel from 190 to 208 Hz to increase naturalness of the vowel token (Assmann & Katz, 2000). In order to remain consistent the duration of this token was also 400 ms. The complete codes and parameters used to create synthetic speech tokens are presented in Appendix G. Male and female synthetic speech tokens were then normalized using Peak 4 software and saved as a .WAV files. The .WAV files were then imported into Cool Edit Version 96 and saved as PC compatible files. Once saved, these tokens were imported into the Stim2 Sound Editor module and bilaterally presented to the participant via insert earphones at a rate of 1.1/s and an intensity of 75 dB pSPL.

Recorded self-produced vowel tokens. Prior to collection, participants were asked to produce the vowel /a/ 10 times into a Logitech (Model 980186-0403) desktop microphone. These tokens were recorded using Bias Sound Creative Peak 4.13 running on a Macintosh PowerPC G4 at a sampling rate of 44,100 Hz. Participants were asked to produce the vowel tokens using a normal vocal effort and speech rate. Complete instructions for vowel recordings are included in Appendix H. The duration of each vowel was then measured. To remain consistent with the other vowel tokens presented in this study the sample token approximately 400 ms was selected. Vowel duration was determined as described above. The average duration of the self-produced male tokens was 388 ms (*SD* = 17.5) and 386 (*SD* = 26.8) for the self-produced female tokens. These tokens were then amplitude normalized and saved as a .WAV file. The .WAV file was then exported and opened using Syntrillium Software Corporation Cool Edit version 96 and

resaved as a PC compatible .WAV files. This step was necessary to ensure that the recorded vowel tokens were compatible with the Compumedics Sound Editor module.

Frequency shifted self-produced vowel tokens. Using Bias Sound Creative Peak 4.13 software, the recorded self-produced vowels were frequency shifted down one-half an octave. These tokens were then amplitude normalized and saved as a .WAV file. The .WAV file was then exported and opened using Syntrillium Software Corporation Cool Edit version 96 and resaved as a PC compatible .WAV file. This step was necessary to ensure that the vowel tokens recorded and saved in Peak 4.13 software were compatible with the Compumedics Sound Editor module.

Tonal stimulus. The center frequency of the tone burst was selected based on methodology presented by Tiitinen et al., (1999). In this study, the center frequency of the tonal stimulus was generated to match the mean center frequency of the highest peak observed in vowels /a/ and /A/. Following this rationale, male and female synthetic /a/ tokens were analyzed using *SpectraPRO*-FFT Spectral Analysis System software Version 3.32 on a Dell Inspiron laptop. The center of the highest harmonic of both the synthetic male and female tokens was 722.66 Hz.

Accordingly, the tone burst was generated with a center frequency of 723 Hz using a Dell Optiplex GX620 computer (Intel (R) Pentium(R) 4 CPU, 2.99 GHz, 1.00 GB RAM, with a 16-bit integrated Creative SB Audigy 4 WDM sound card operating on a Microsoft Windows XP operating system) and Compumedics Stim2 Sound Editor2 software at a sampling rate of 48,000 Hz. A Blackman window was employed with a 10.5 ms ramp and a 49 ms plateau. The total duration was 70 ms.

Acoustic Analysis of Experimental Stimuli

The amplitude as a function of time waveform for the tone burst is presented Figure 2 and for the speech tokens including the natural tokens, synthetic tokens, an example of one male and one female self-recorded, and the self-recorded frequency shifted tokens are presented in Figures 3 through 10. Waveforms were initially generated using *SpectraPRO*-FFT Spectral Analysis System Version 3.32 software. Data points were then copied as text files and saved using Microsoft Notepad. These points were then imported into Excel files, saved, and exported into Cricket Graph III Version 1.5.3. Cricket Graph was used to generate graphs.

Fast Fourier Transforms (FFTs) were performed on these tokens using *SpectraPRO*-FFT Spectral Analysis System Version 3.32 software (see Figures 11 through 19). The four settings used to configure the FFTs were the smoothing window, the sampling rate (Hz), the FFT size, and the decimation ratio. The purpose of applying a smoothing window is to reduce the amount of spectral leakage or the amount of "energy in the signal "leaks" from its proper location into the adjacent lines" (*SpectraPRO* User Guide, p. 37). FFT sampling rate refers to the number of times per second an analog signal is sampled or digitized to construct the digital representation of the original signal. When analyzing a prerecorded-preconstructed sound file (i.e., a post recording analysis) the sampling rate for the FFT is determined by the sampling rate of the file and cannot be changed. The values presented below were pre-determined and unchangeable according to how the file was originally recorded. When using *SpectraPRO* the FFT size is related to the frequency resolution. The larger the FFT size the higher the frequency resolution of the FFT. FFT size is also related to the number of spectral lines or points. The number of points is one-half the FFT size. For example, there will be 512 points or spectral lines in an analysis with a FFT size of 1024. The decimation ratio refers to the ratio in which the file will be

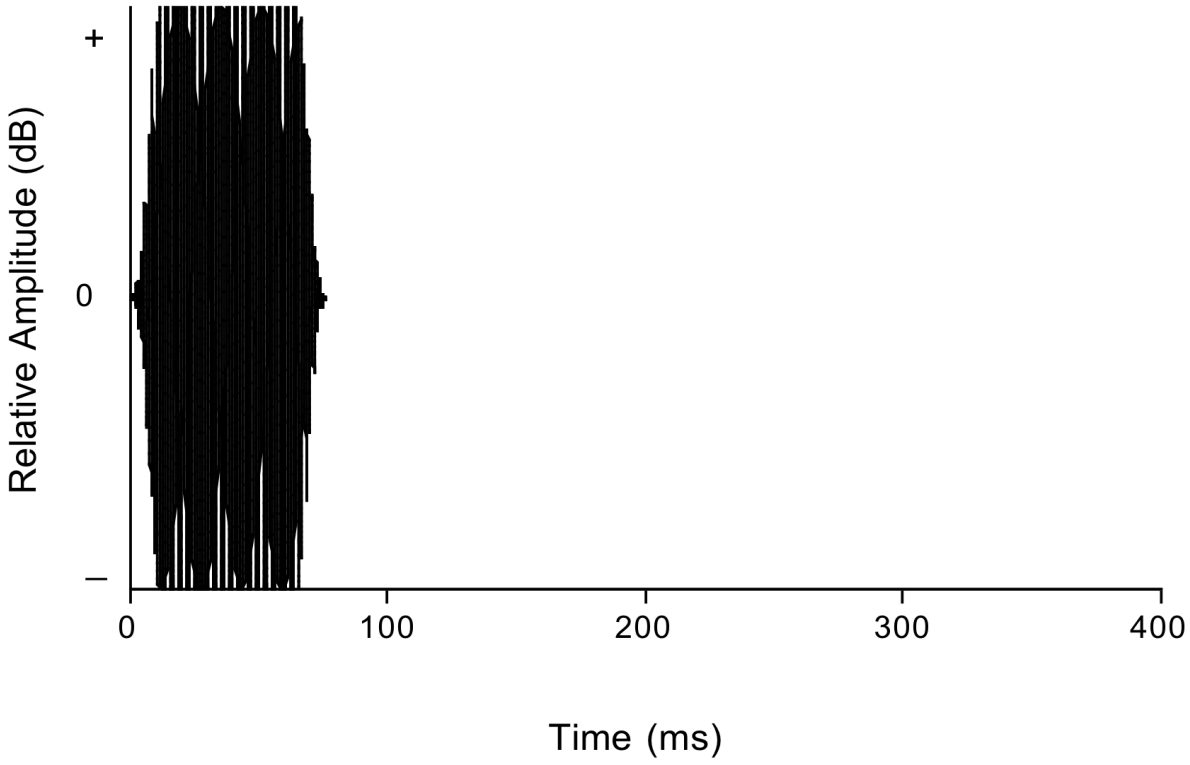


Figure 2. Amplitude as a function of time for the tonal stimulus.

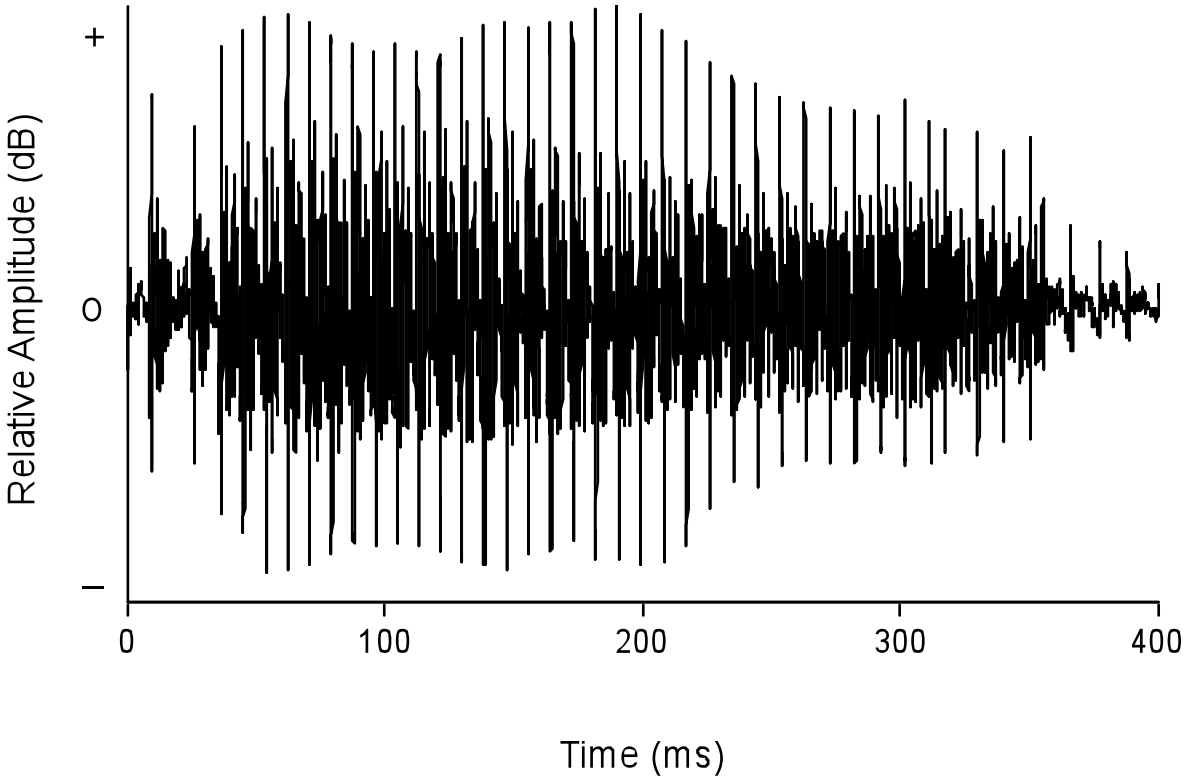


Figure 3. Amplitude as a function of time for the natural male /a/ stimulus token.

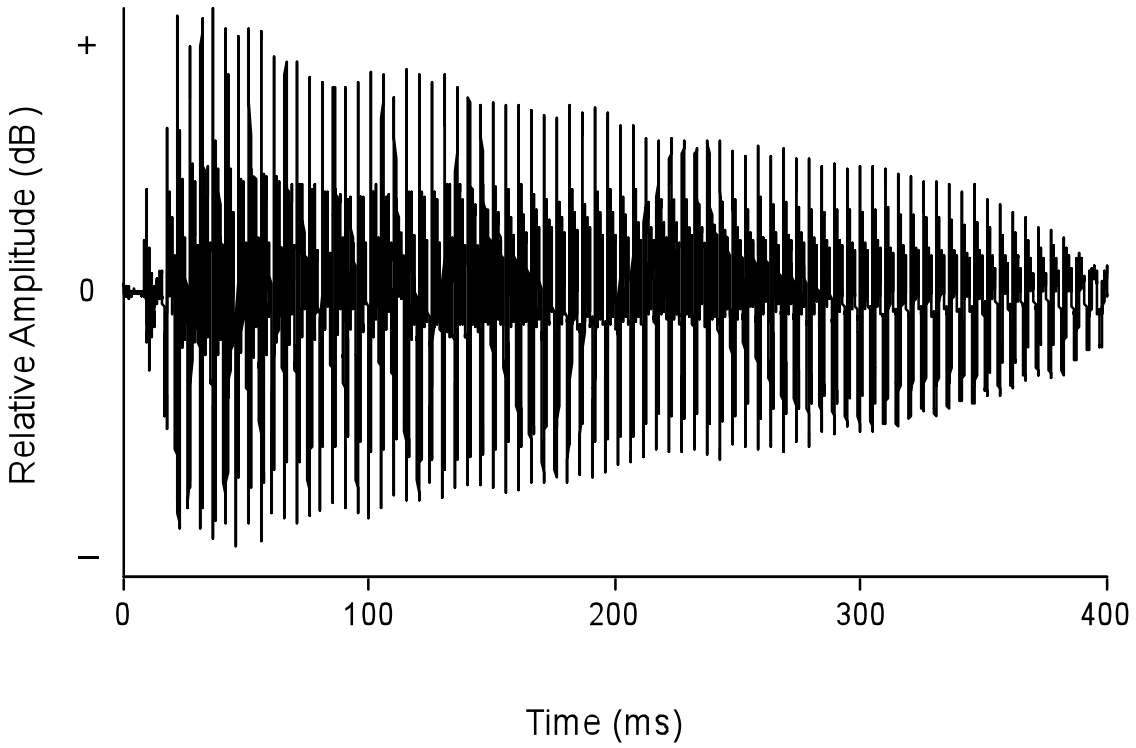


Figure 4. Amplitude as a function of time for the natural female /a/ token.

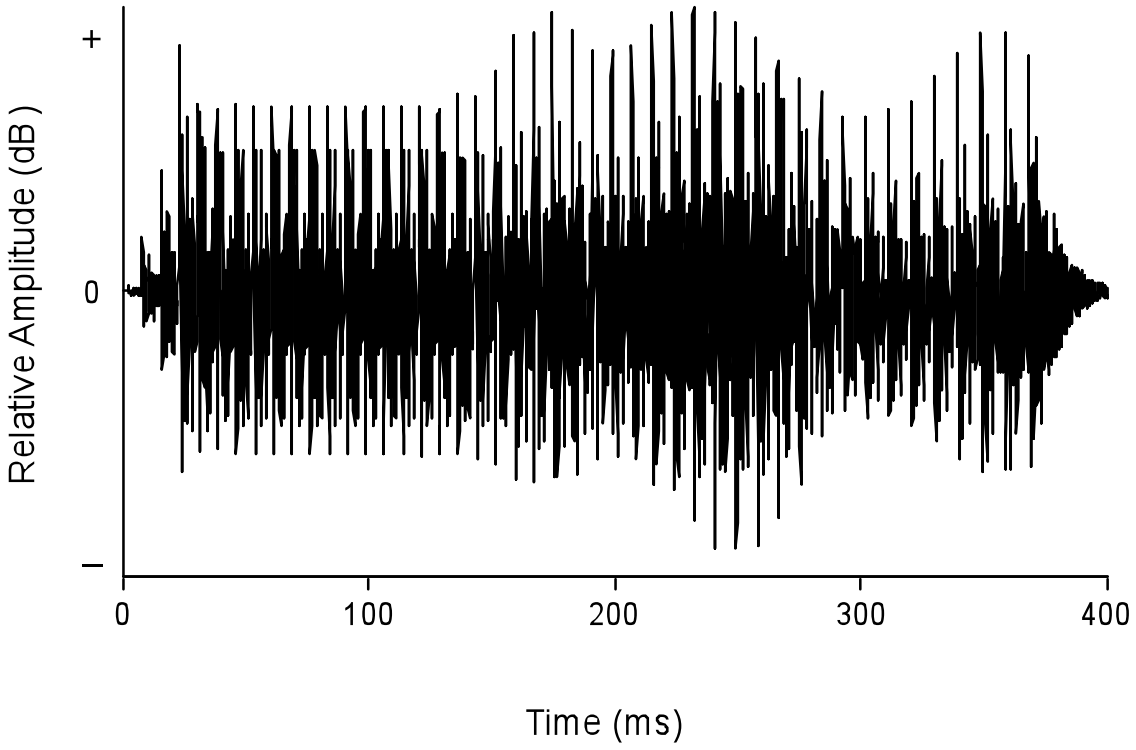


Figure 5. Amplitude as a function of time for the synthetic male /a/ token.

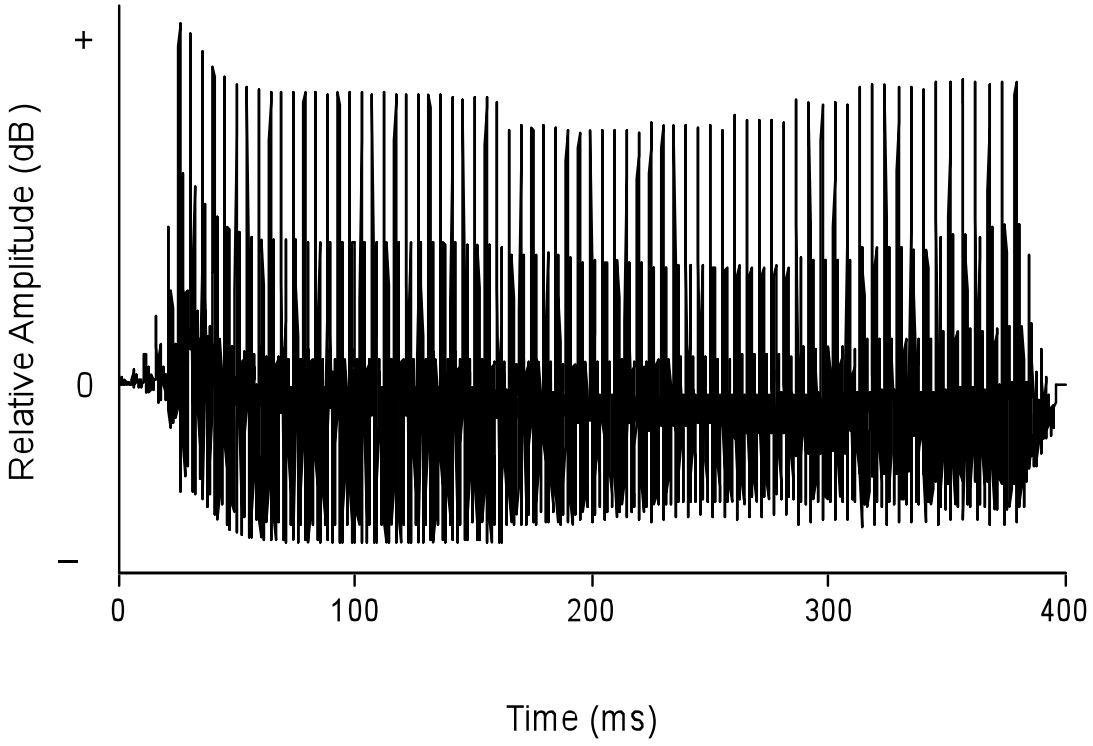


Figure 6. Amplitude as a function of time for the synthetic female /a/ token.

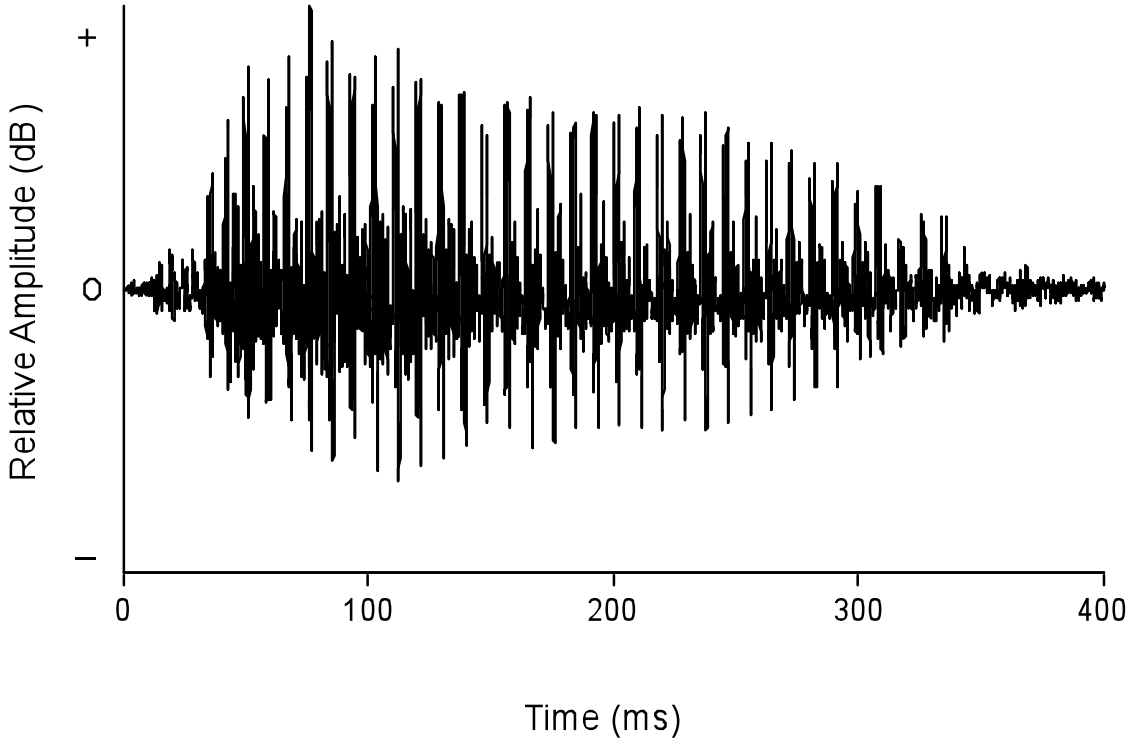


Figure 7. Amplitude as a function of time for an example of one male self-produced /a/ token.

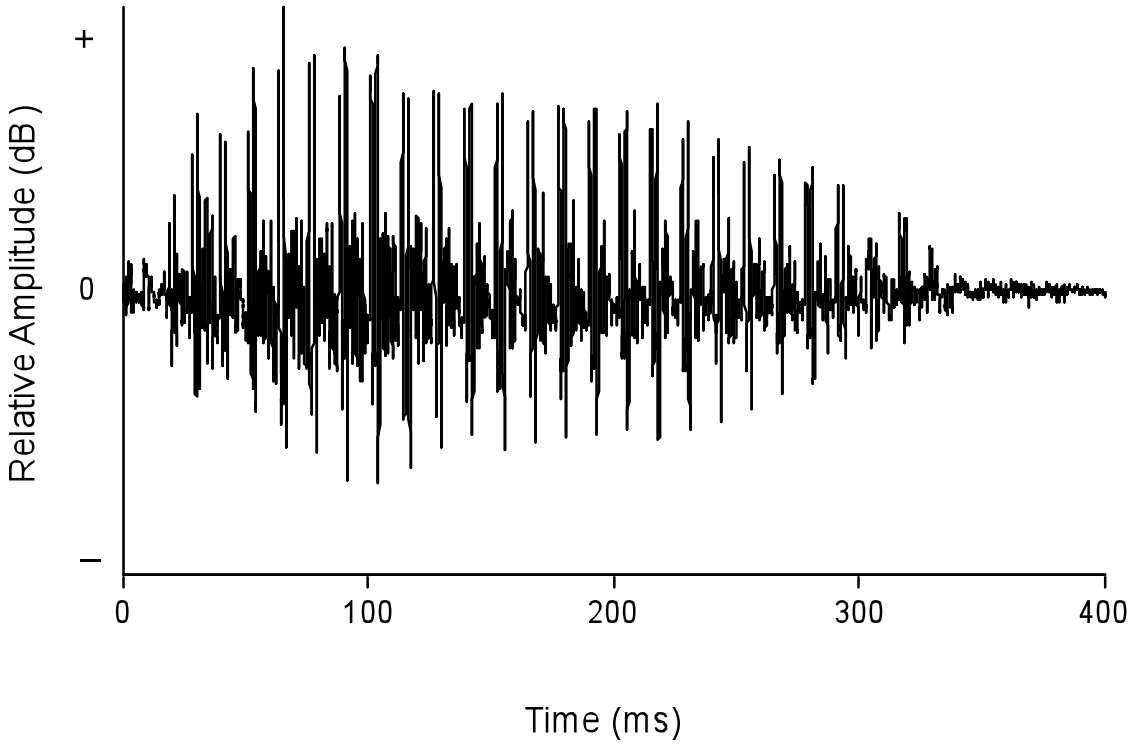


Figure 8. Amplitude as a function of time for an example of one male frequency shifted self-produced /a/ token.

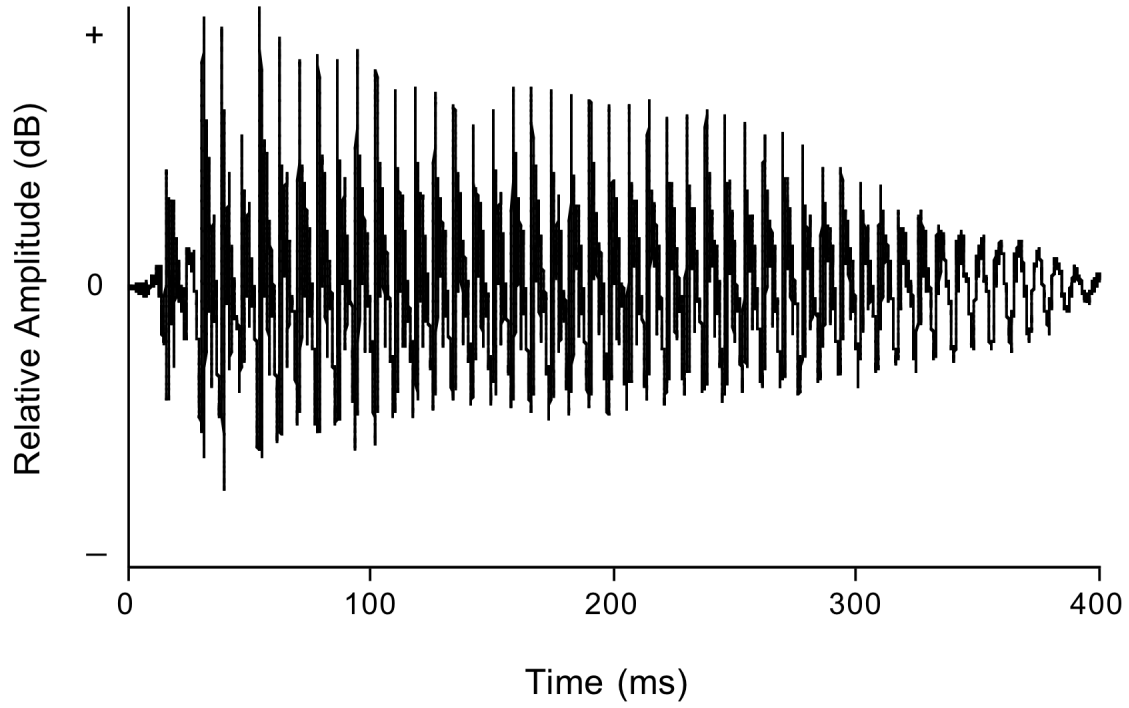


Figure 9. Amplitude as a function of time for an example of one female self-produced /a/ token.

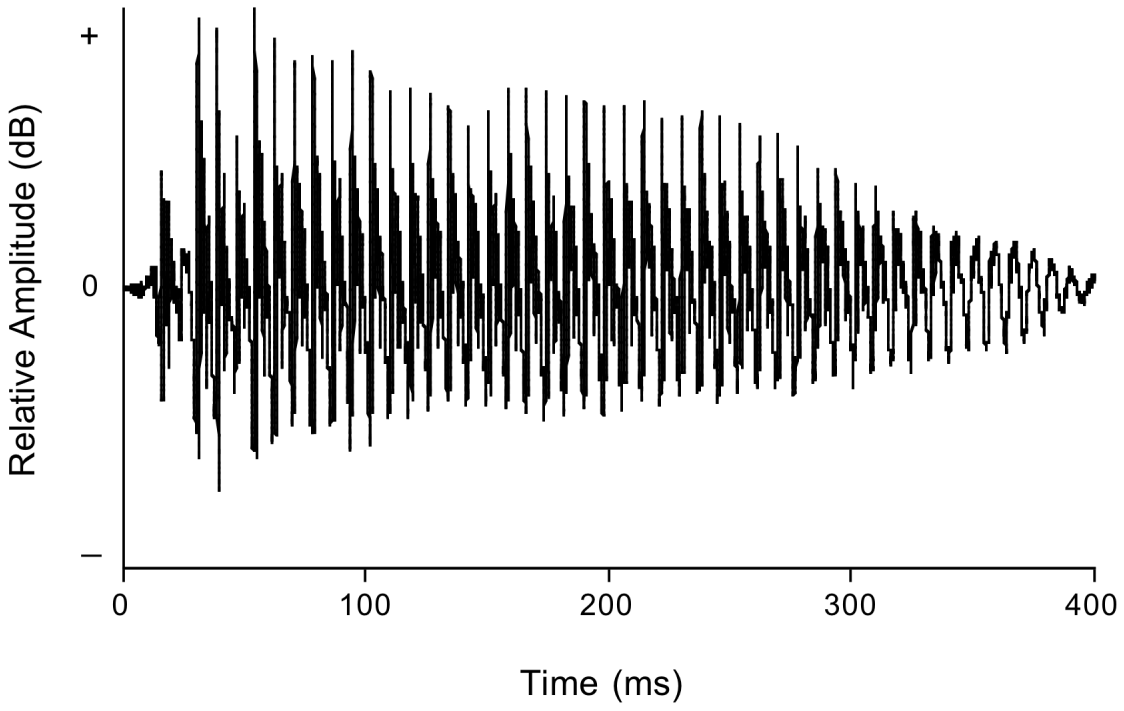


Figure 10. Amplitude as a function of time for an example of one female frequency shifted self-produced /a/ token.

"down sampled" or reproduced with a lower sampling rate. If the decimation ratio were set at a ratio of 8:1, the *SpectraPRO* software would average eight samples together and produce one sample. A decimation ratio of 1 does not change or down sample the original signal. The specific FFT settings for the tokens used in the present investigation are presented following.

The FFT for the tone was generated using a Hanning window with a sampling rate of 48, 000 Hz, a FFT size of 2048 Hz, and a decimation ratio of 1. FFTs of both male and female natural vowel tokens were generated using a Hanning window with a sampling rate of 44, 100 Hz, a FFT size of 2048, and a decimation ratio of 1.

FFTs of the synthetic male and female vowel tokens were generated using a Hanning window a 10,000 Hz sampling rate, a FFT size of 1024, and a decimation ratio of 1. Data points from the FFTs were then copied as text files and saved using Microsoft Notepad. These points were then imported into Excel files, saved, and exported into Cricket Graph III Version 1.5.3. Cricket Graph was used to generate graphs. Ideally, all FFT graphs would have been constructed using identical parameters (i.e., sampling rate and size); however, this was not possible due to the differences in the settings used to initially record and construct the individual tokens. Again, when performing a post recording analysis the sampling rates of the recorded file govern the FFT sampling rate and cannot be manipulated.

Spectrograms for the speech tokens were also constructed (see Figures 20 to 27). All spectrograms were generated using Praat Version 5.0.32 software (Boersma & Weenink, 2008) retrieved from http://www.praat.org. Stimuli .WAV files were "Read" (i.e., opened) in the Praat software and analyzed using the "To Spectrogram" software function. A Hanning "window shape" which defines the shape of the analysis window (Boersma & Weenink, 2008) was selected. Default parameters were maintained for all the other window settings and the

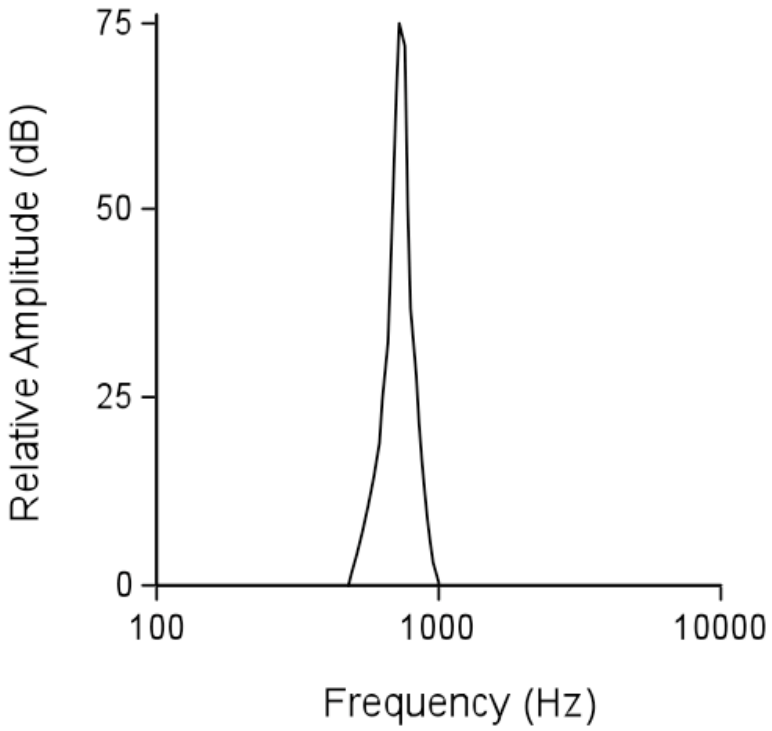


Figure 11. FFT of the experimental tone burst.

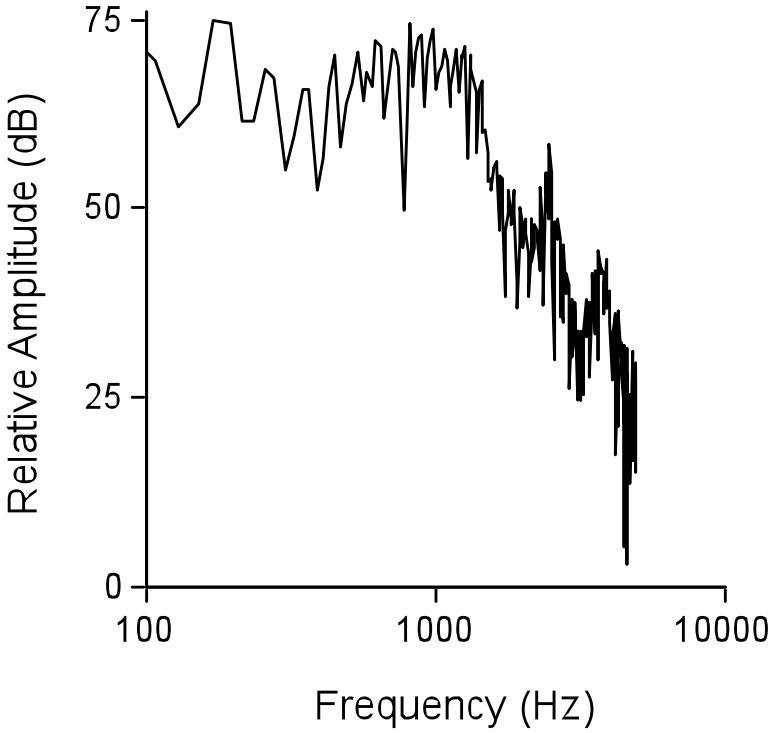


Figure 12. FFT of the experimental natural male /a/ token.

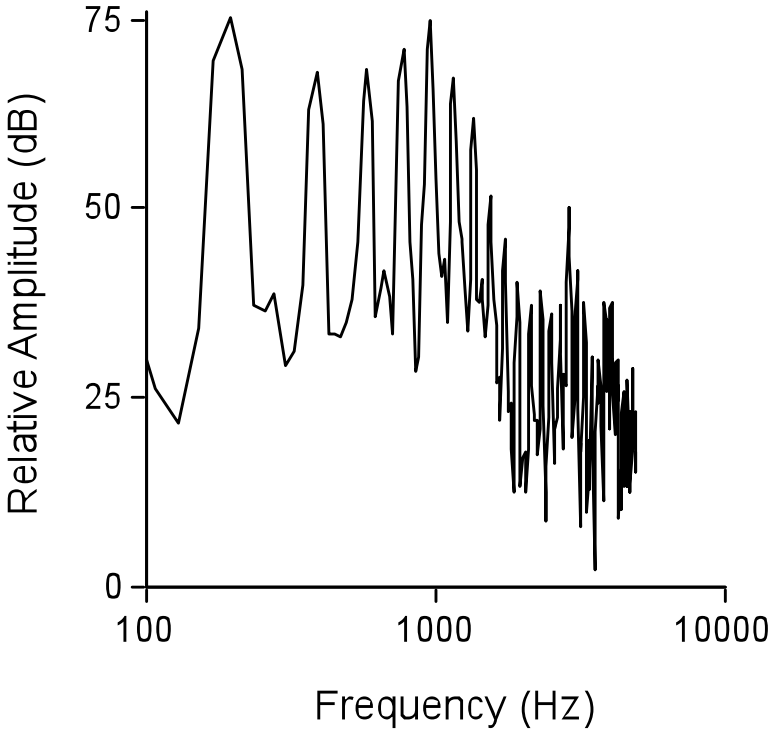


Figure 13. FFT of the experimental natural female /a/ token.

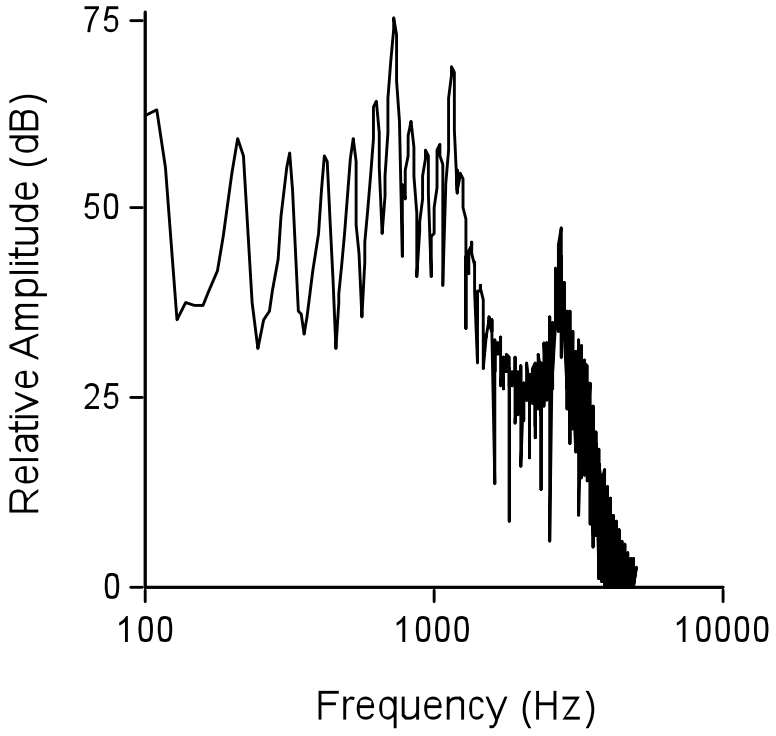


Figure 14. FFT of the experimental synthetic male /a/ token.

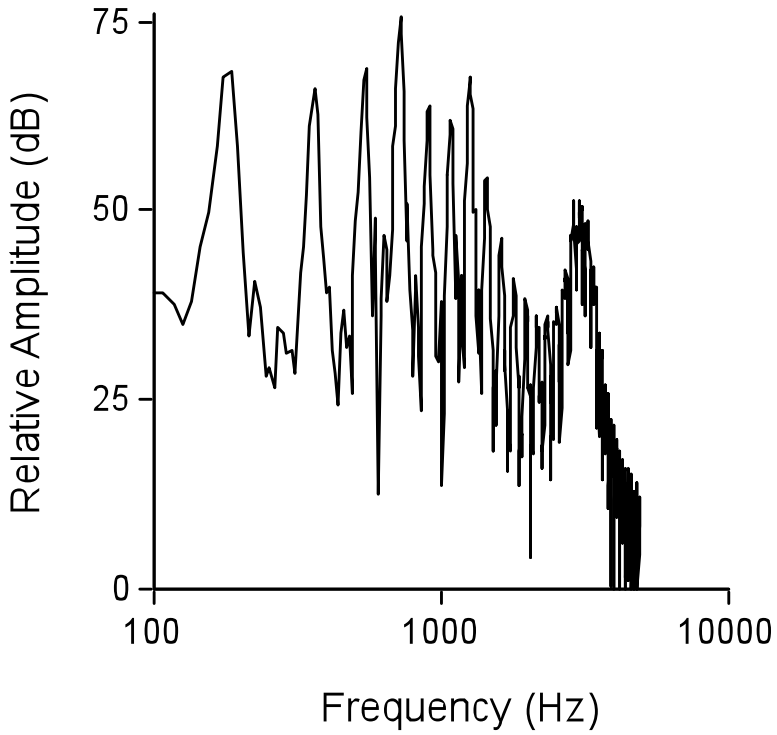


Figure 15. FFT of the experimental synthetic female /a/ token.

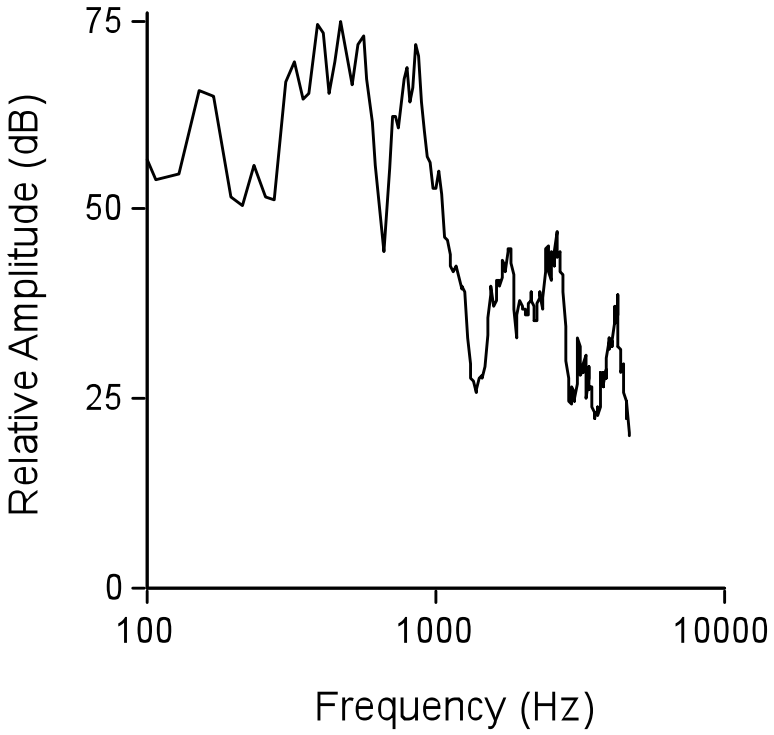


Figure 16. FFT of the male self-produced /a/ token example.

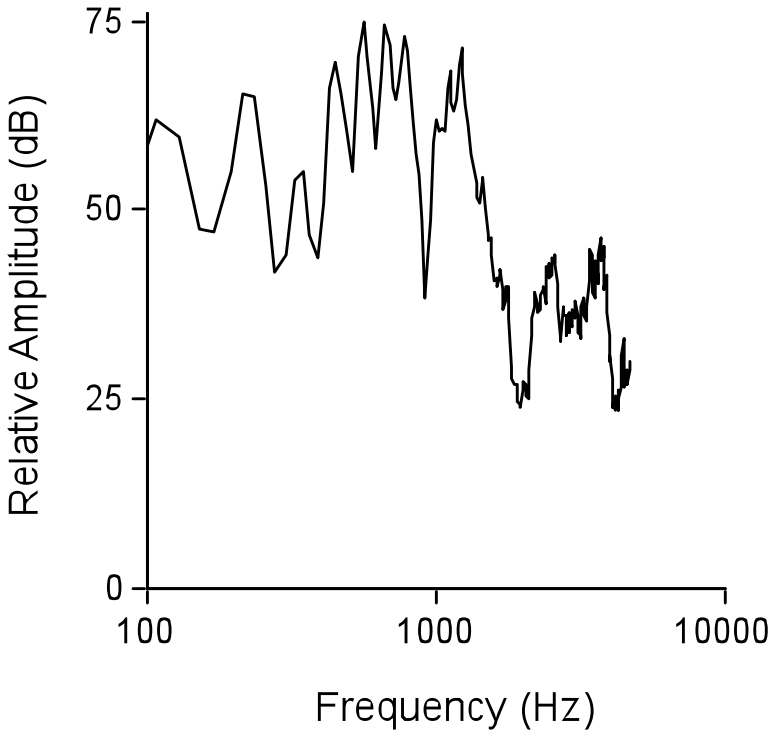


Figure 17. FFT of the male frequency shifted self-produced /a/ token example.

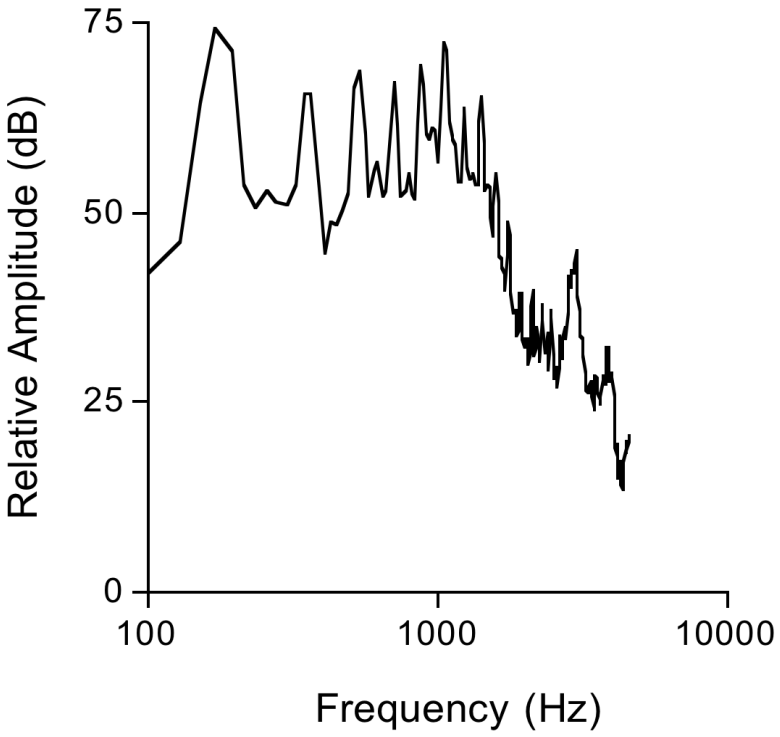


Figure 18. FFT of the female self-produced /a/ token example.

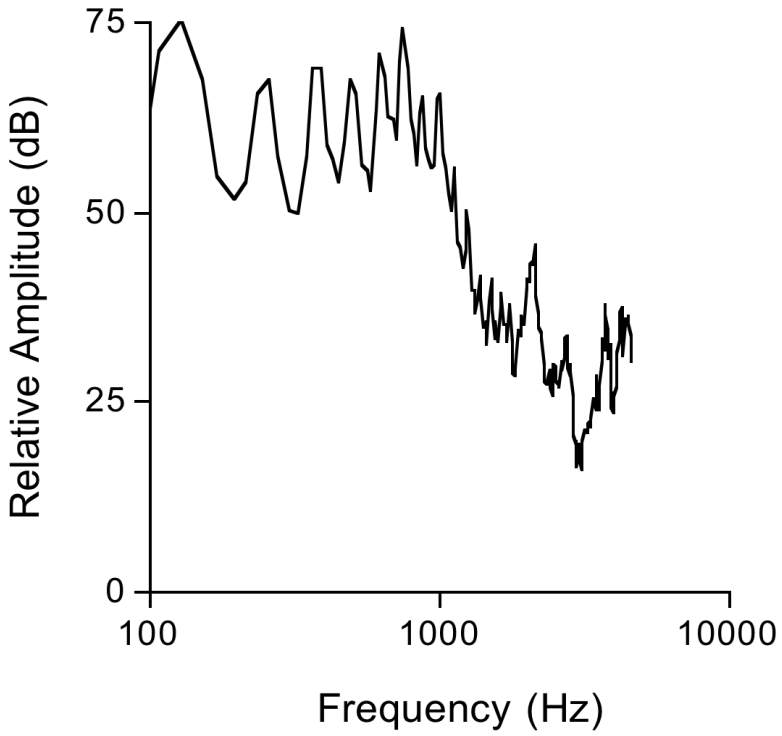


Figure 19. FFT of the female frequency shifted self-produced /a/ token example.

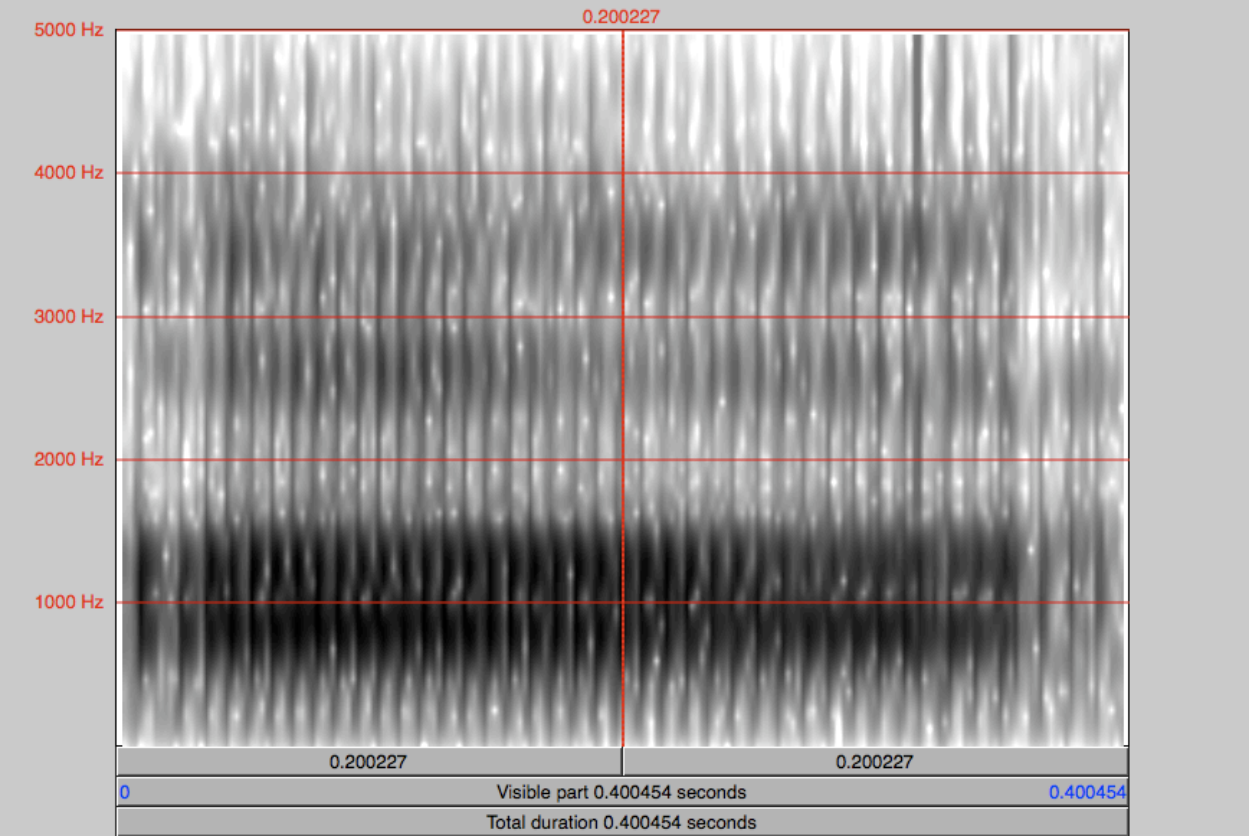


Figure 20. Spectrogram of natural male /a/ token.

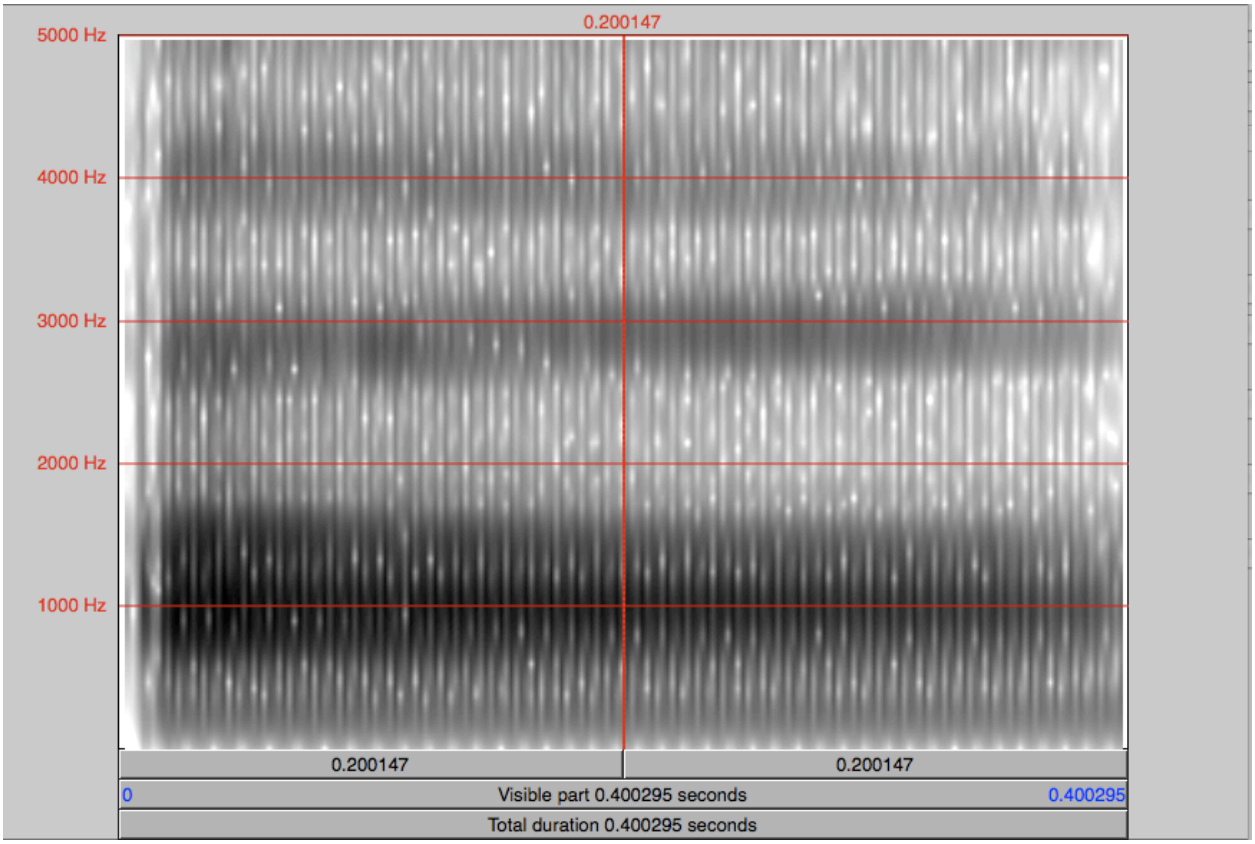


Figure 21. Spectrogram of natural female /a/ token.

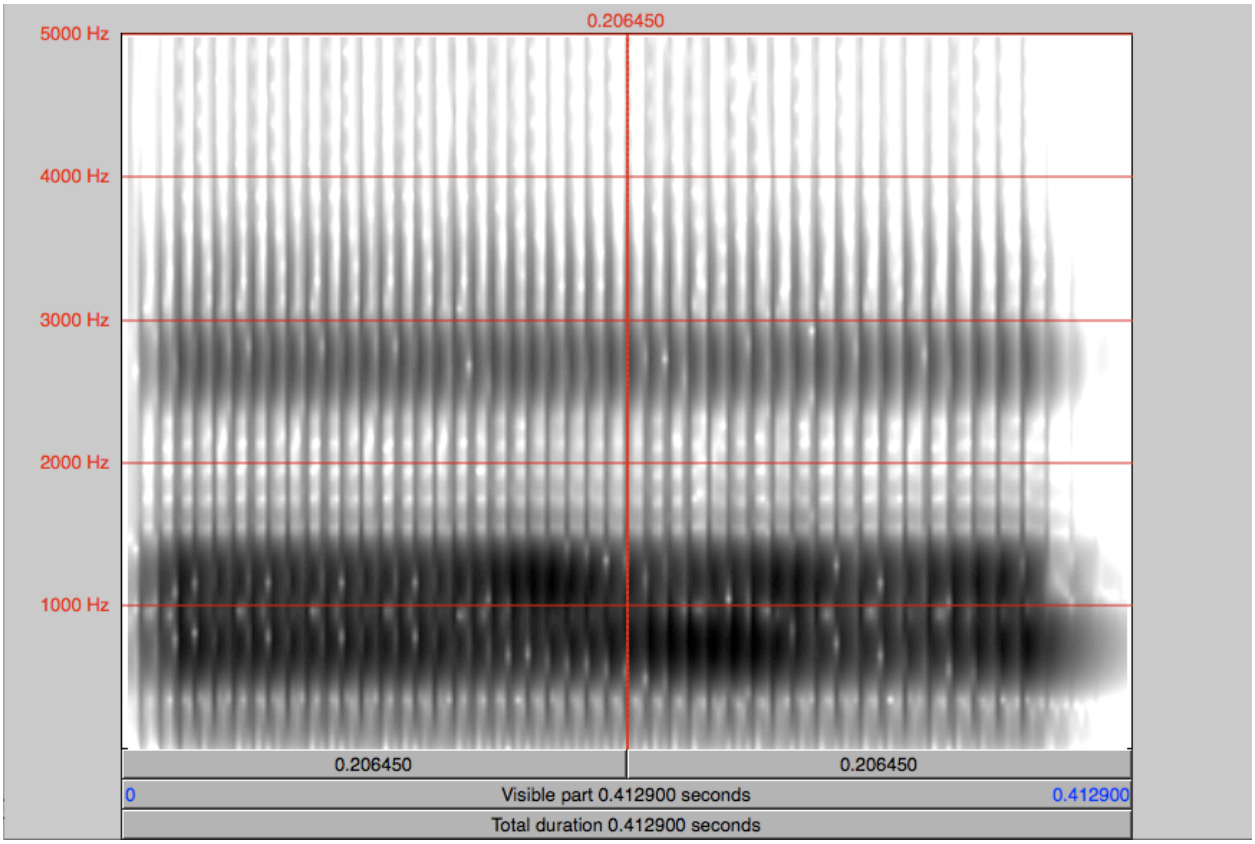


Figure 22. Spectrogram of synthetic male /a/ token.

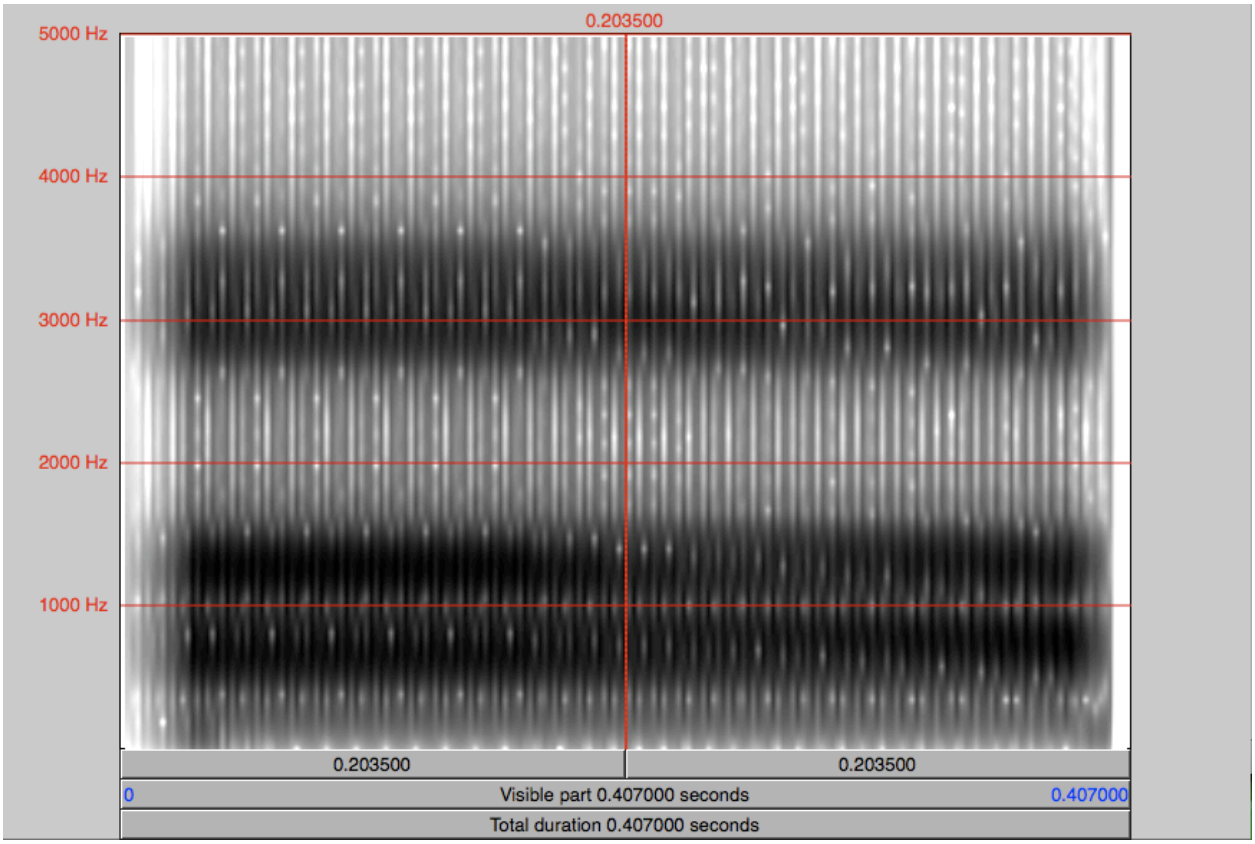


Figure 23. Spectrogram of synthetic female /a/.

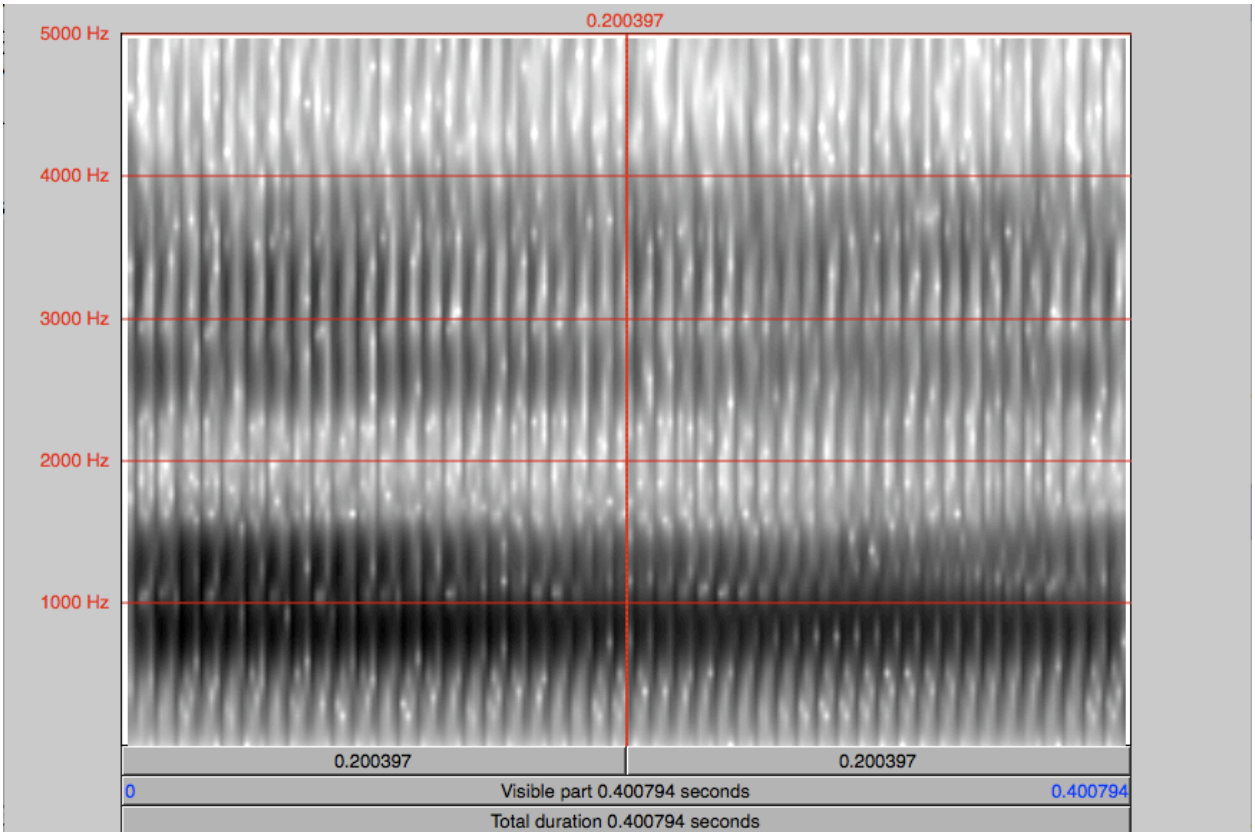


Figure 24. Spectrogram of male self-produced /a/ token example.

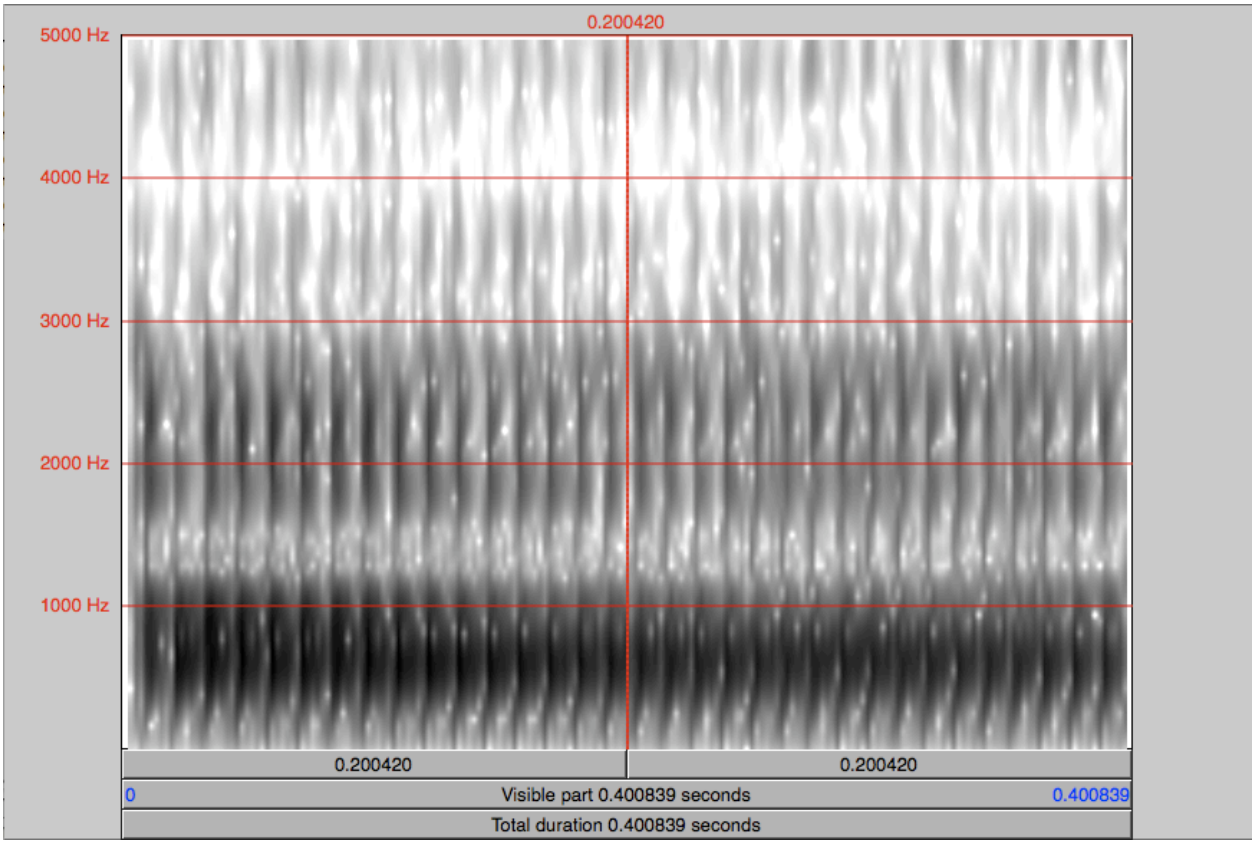


Figure 25. Spectrogram of male frequency shifted self-produced /a/ token example.

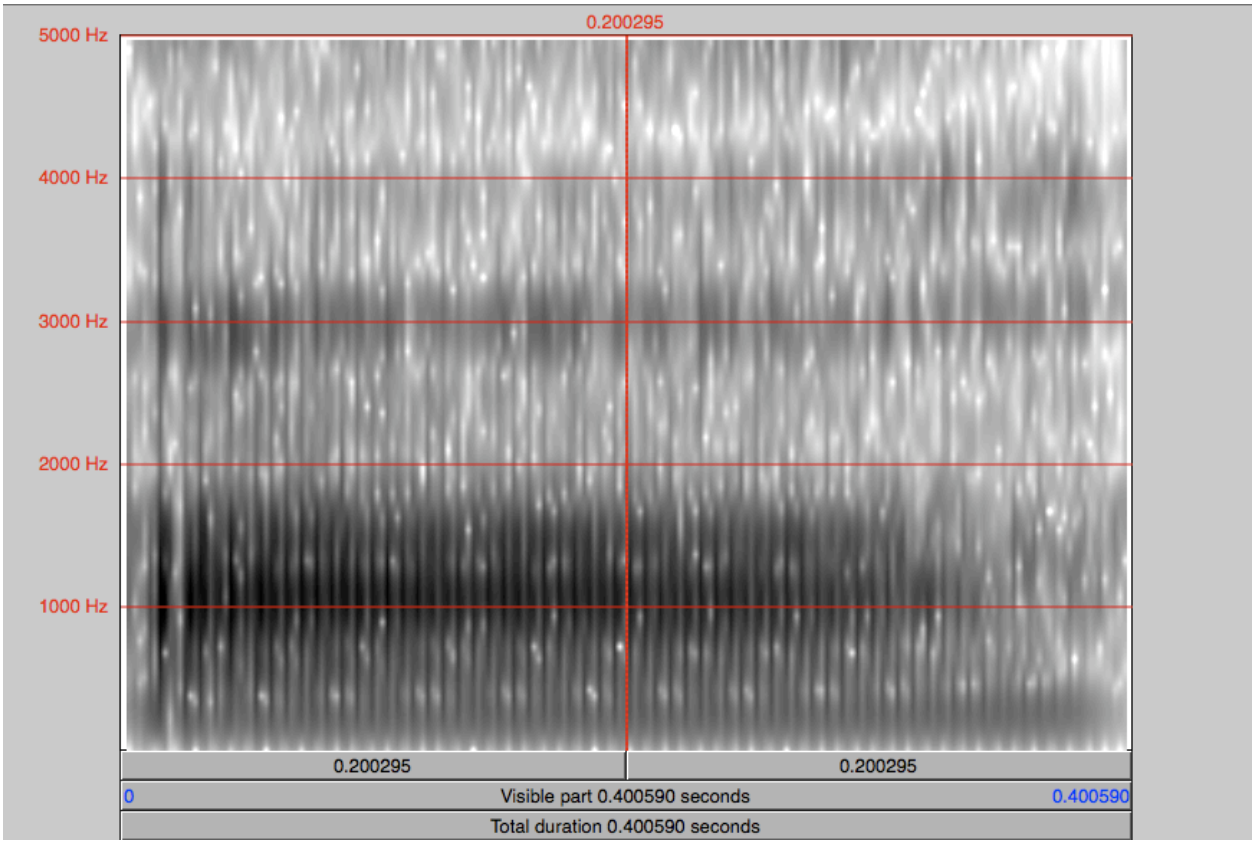


Figure 26. Spectrogram of female self-produced /a/ token example.

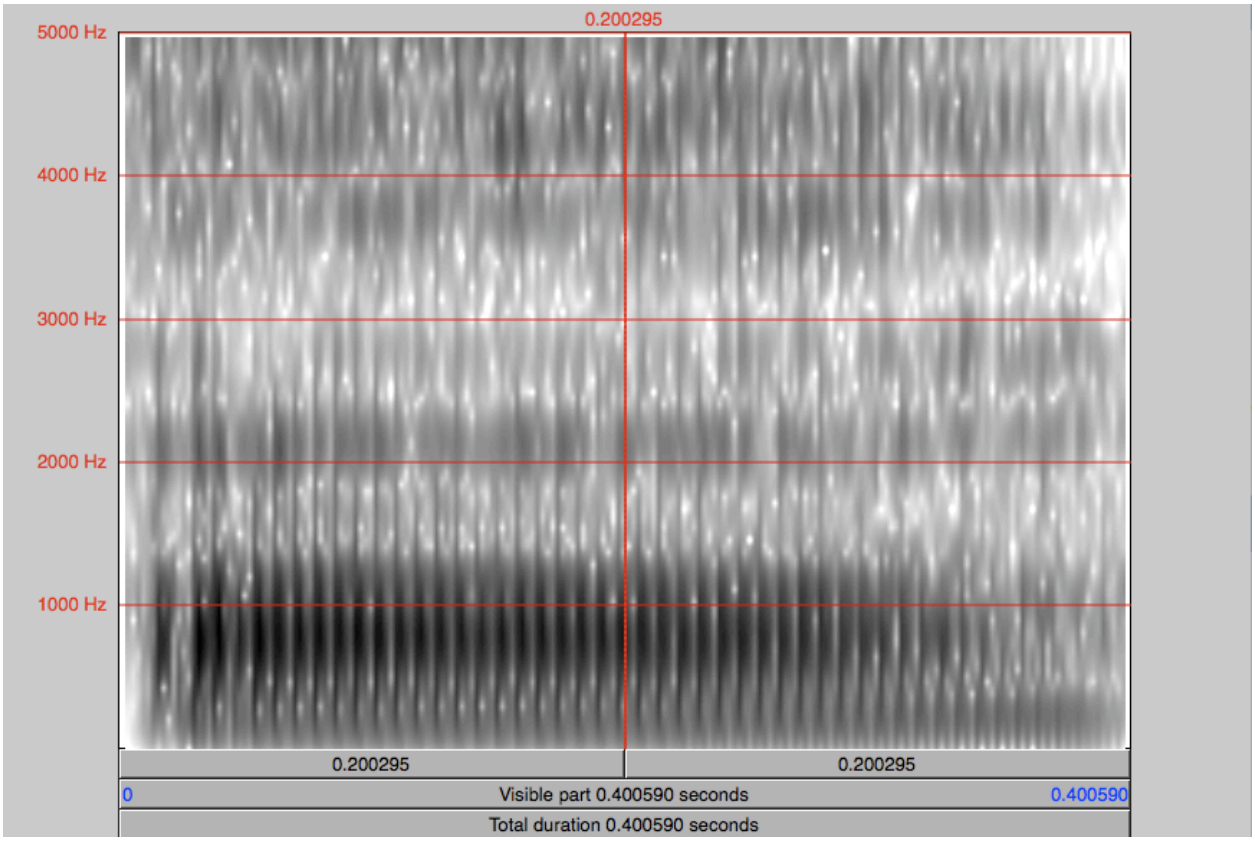


Figure 27. Spectrogram of female frequency shifted self-produced /a/ token example.

corresponding spectrogram formulated.

Stimulus Calibration

All stimulus presentation levels were calibrated to 75 dB pSP L using a Brüel & Kjær precision sound level meter (type 2231) with a Brüel & Kjær octave band filter (type 1625) attached to a Brüel & Kjær (type 4144) pressure microphone. In order to keep calibration parameters consistent across all tokens, calibration was performed using pSPL measurements.All stimuli were presented using the Sound Editor Module, routed to an insert earphone (NeuroScan ER-3A) that was coupled to a 2cc (Brüel & Kjær type DB 0138) coupler attached to the sound level meter.

Procedure

Prior to testing, the circumference of the participant's head was measured to ensure appropriate cap sizing. The Compumedics NeuroScan Quik-Caps come in various sizes and styles. Three sizes were available for use during this experiment. Head circumference measurements of 50 to 55 cm, 56 to 59 cm, and 60 to 65 cm fit cap sizes small, medium, and large, respectively. Electrodes are placed on the Quik-Cap according to the 10/20 International Electrode System (Jasper, 1958).

Stimuli presentations were counter balanced according to a diagram-balance Latin squares design (Wagenaar, 1969) and presented to the listener via NeuroScan ER-3A insert earphones. Insert earphones were placed according to methodology presented by Dean and Martin (2000) for deep insert earphone placement. The procedure used was as follows: the insert eartip was squeezed between the fingertips, initially placed into the participant's ear canal, once iplaced, the pinna was pulled back and the eartip placed so that the outer part is even with the canal (Dean & Martin, 2000). All stimuli were presented bilaterally at a rate of 1.1/s.

During the experiment all participants were asked to sit quietly and silently count the number of presented tokens. All were asked to write the number of tokens they counted for each condition on the sheet of paper provided. They were also asked to remain alert and minimize movements, especially head and neck movements, while the stimuli were presented. Frequent breaks were given throughout testing to ensure participant comfort and avoid listener fatigue. During breaks, the insert earphones were removed and the Quik-Cap was disconnected from the NeuroScan electrode headbox. Disconnecting the cap from the headbox allowed the participant to be able to move freely while still wearing the cap. Once the participant was ready to continue testing, cap placement on the participant's head was evaluated to ensure that the cap had not shifted, the cap was reconnected to the headbox, and the insert earphones were replaced in the participant's ears. Electrode impedances were re-checked, if needed, additional electrode gel was added to individual electrode sites. Testing began once all electrode sites maintained impedances of $\leq 5000\ \Omega$.

EEG activity was recorded from eleven electrode sites placed about the participant's head (i.e., F3, Fz, F4, C3, Cz, C4, T4, T3, Pz, M1, and M2). The nasion served as the reference and Fpz as the common ground. Vertical eye movements were monitored using electrodes placed vertically above and below the left eye. Electrode impedances were maintained at or below 5000 Ω. EEG activity was amplified 3000 times and analog filtered (i.e., 1-30 Hz). Analog-to-digital conversion at a rate of 500/s for all channels was performed with a PC based NeuroScan system and SynAmps^2 24 bit amplifier. All EEG activity was saved as continuous files for offline averaging and digital filtering.

Off-line Waveform Analyses

All continuous files were epoched in a time locked window from -100 to +500 ms

relative to stimulus onset. This response window was selected to provide an adequate time window for visualization of all pertinent waveform components (Martin & Stapells, 2005; Sharma, Marsh, & Dorman, 2000). A pre-stimulus interval of -100 was used as the baseline correction for all recording channels. Epochs were then digitally filtered (i.e., 1-40 Hz, 24 dB/octave) and re-referenced off-line to linked mastoids (M1 and M2). Ocular movement artifacts were digitally removed from the epochs (Semlitsch, Anderer, Schuster, & Presslich, 1986). Epochs containing artifacts exceeding +/- 50 μV were rejected from averaging. All waveforms components were replicated. Replication was defined as two or more waveforms with identifiable P1-N1 and/or P2 peaks within 25 ms (O'Brien & Stuart, 2001). Both replications were then averaged creating seven passive waveforms for each individual participant.

Electrophysiological Waveform Analyses

Waveform components P1-N1-P2 were evaluated in terms of latency and amplitude measures. P1 was defined as the largest positive peak following stimulus onset between 40 and 150 ms. N1 was defined as the largest negative deflection following P1 between 75 and 210 ms and P2 as the following positive peak between N1 and 300 ms. Based on values used by Kraus et al. (1993) P1-N1-P2 wave components were considered present if amplitudes were equal to or exceeded 0.5 μV. Amplitudes were measured from component peak to the adjacent trough. Waveform components were also considered present if greater responses were visualized that the fronto-central electrode sites (i.e., Fz and Cz) in comparison to the parietal electrode (i.e., Pz). Polarity inversion at the mastoid sites was also used to determine response presence (Vaughan & Ritter, 1970). The response, however, was not deemed absent if a polarity inversion did not occur (Martin & Stapells, 2005; Tremblay, Piskosz, & Souza, 2003). Examples of latency and

amplitude measurements are presented in Figures 28 and 29, respectively.

Results

Grand average N1-P2 complex amplitudes were measured at all electrode sites to determine the electrode location where the most robust response was recorded. The grand average waveforms recorded across all electrode sites from the male participants are presented in Figures 30, 34, 38, 42, 46, 50, and 54 and presented for the female participants Figures 31, 35, 39, 43, 47, 51, and 55. For all conditions, the N1-P2 complex amplitude was greatest at the Cz electrode site; therefore, all numerical data presented in the current investigation was collected from the Cz electrode site only. Individual participant waveforms recorded from Cz are presented for each experimental condition in Figures 32, 36, 40, 44, 48, 52, and 56 as a function of amplitude and time. Individual waveforms for male participants are labeled as such and are presented in the graph on the left of each figure and the individual waveforms for the female participants are also labeled and presented in the graph on the right. The grand average waveforms are displayed as the thickest black line. Grand average only waves are presented in Figures 33, 37, 41, 45, 49, 53, and 57.

P1 Latency

P1 waveform component mean latencies and standard deviations, as a function of stimulus condition and gender, are presented in Table 2 and graphically presented in Figure 58. A two-factor mixed analysis of variance (ANOVA) was undertaken to examine the effect of stimulus condition and gender. The ANOVA summary for the P1 wave component is presented in Table 3. As evident in Table 3, only a main effect of condition was found. To investigate the source of the main effect of stimulus condition on the P1 wave component, five orthogonal single-*df* contrasts of interest were undertaken (Keppel & Wickens, 2004). Summaries of those

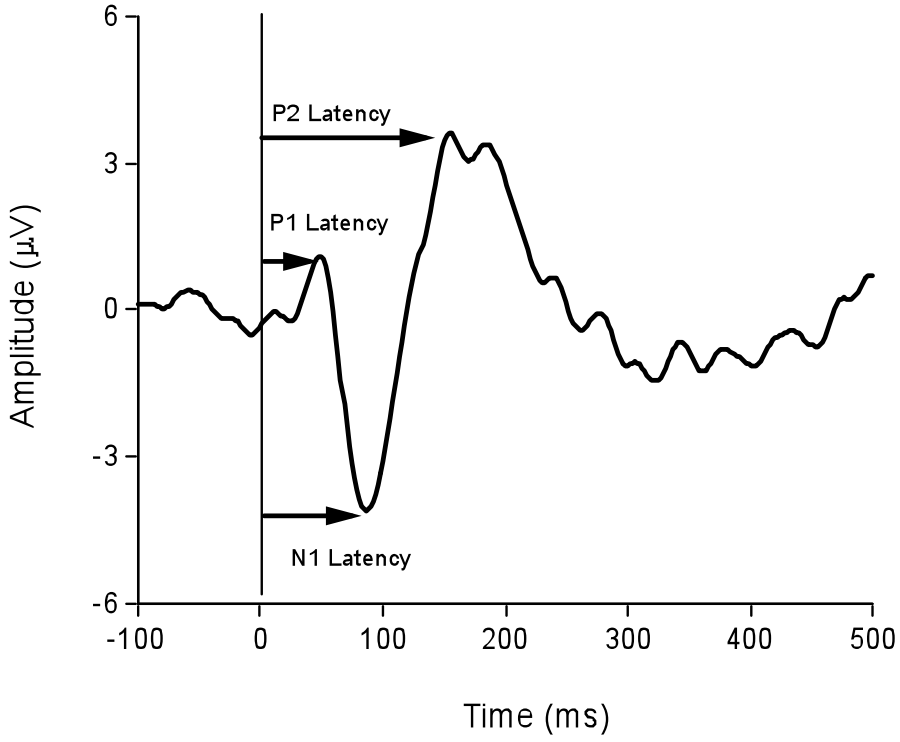


Figure 28. Example of criteria used to determine P1-N1-P2 latency values. Vertical line demonstrates stimulus onset.

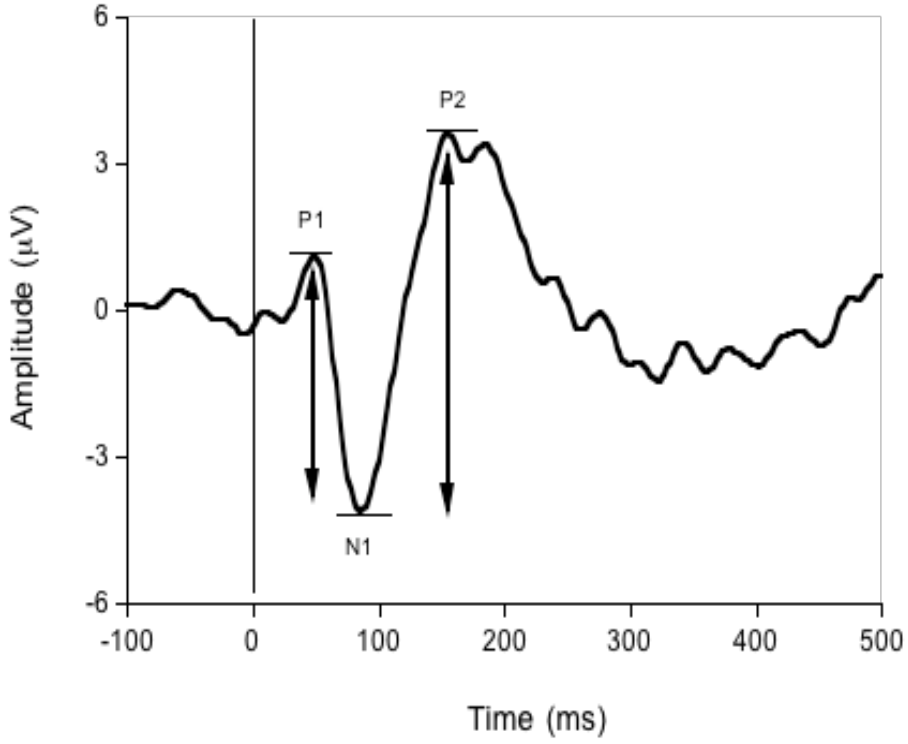


Figure 29. Example of criteria used to determine P1-N1 and N1- P2 amplitude values. Vertical line demonstrates stimulus onset. Horizontal lines represent peak-to-peak amplitudes.

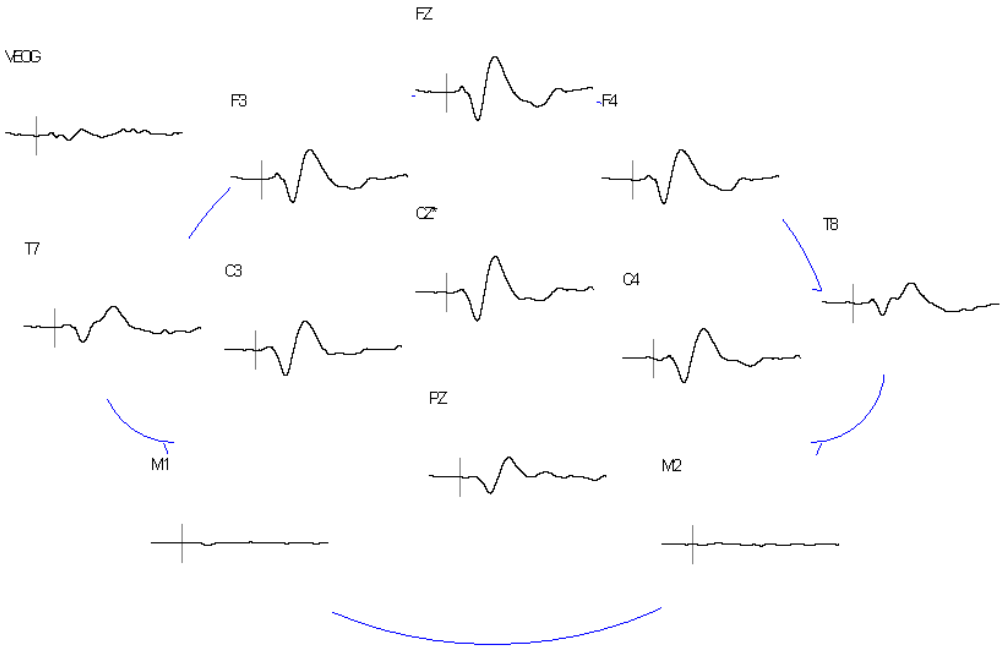


Figure 30. Grand average waveforms recorded across all electrode sites from male participants elicited via tone bursts.

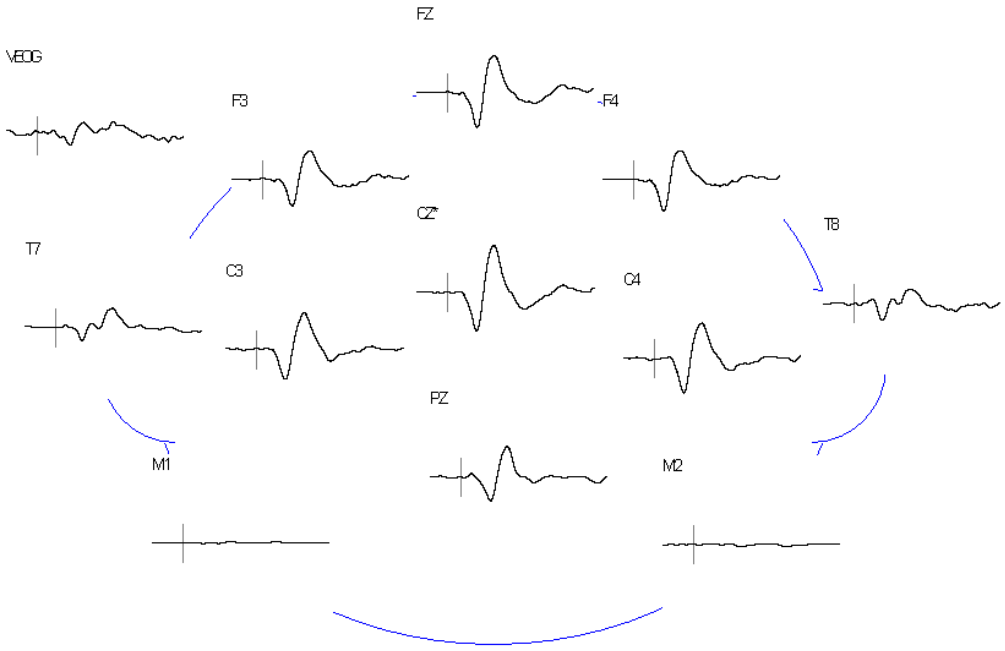


Figure 31. Grand average waveforms recorded across all electrode sites from female participants elicited via tone bursts.

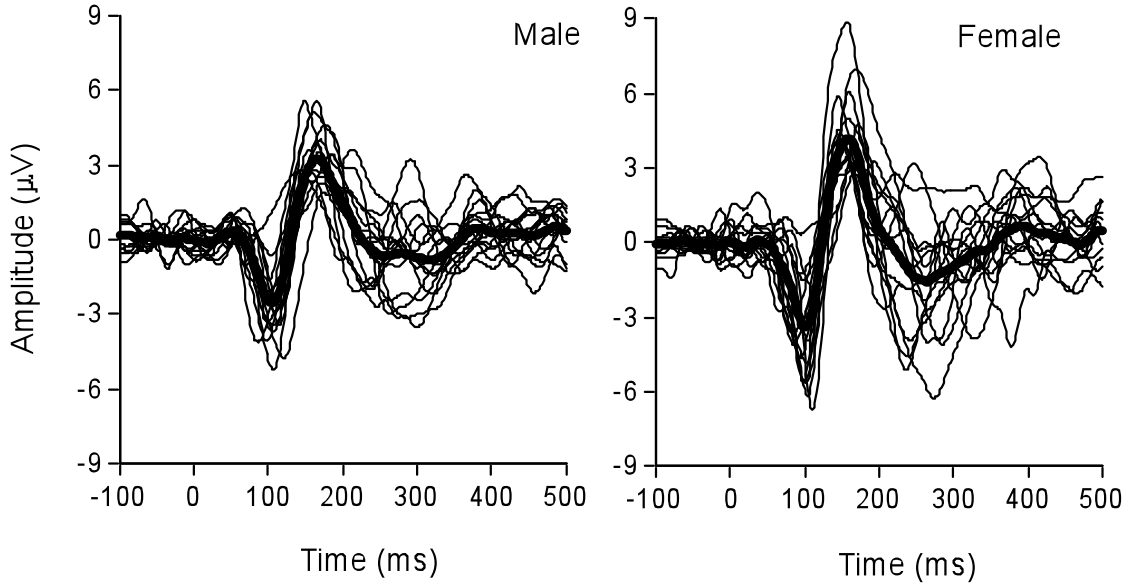


Figure 32. Individual participant waveforms recorded at Cz from male and female participants elicited via tone bursts as a function of amplitude (μV) and time (ms). The grand average waveforms are displayed as the thickest black line.

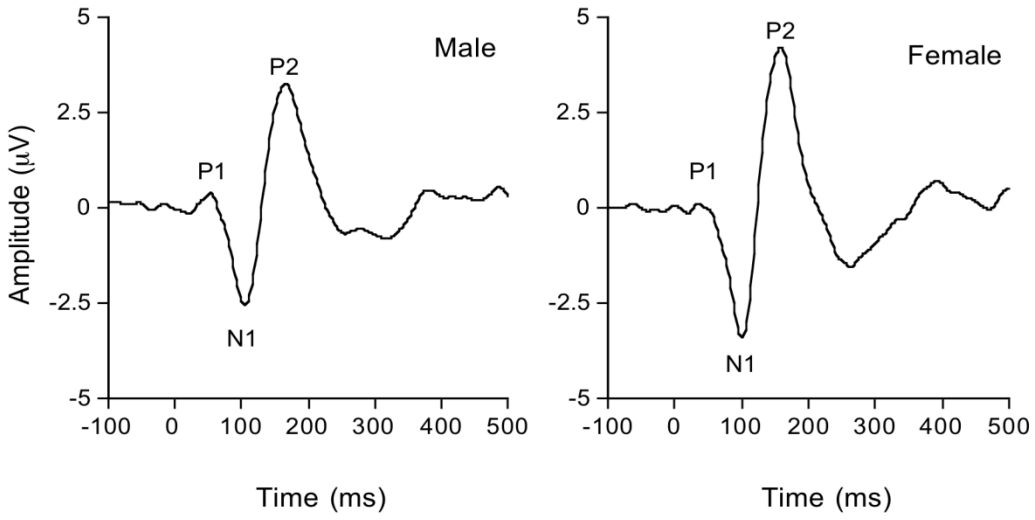


Figure 33. Grand average waveforms recorded at Cz for male and female participants elicited via tone bursts as a function of amplitude (μV) and time (ms).

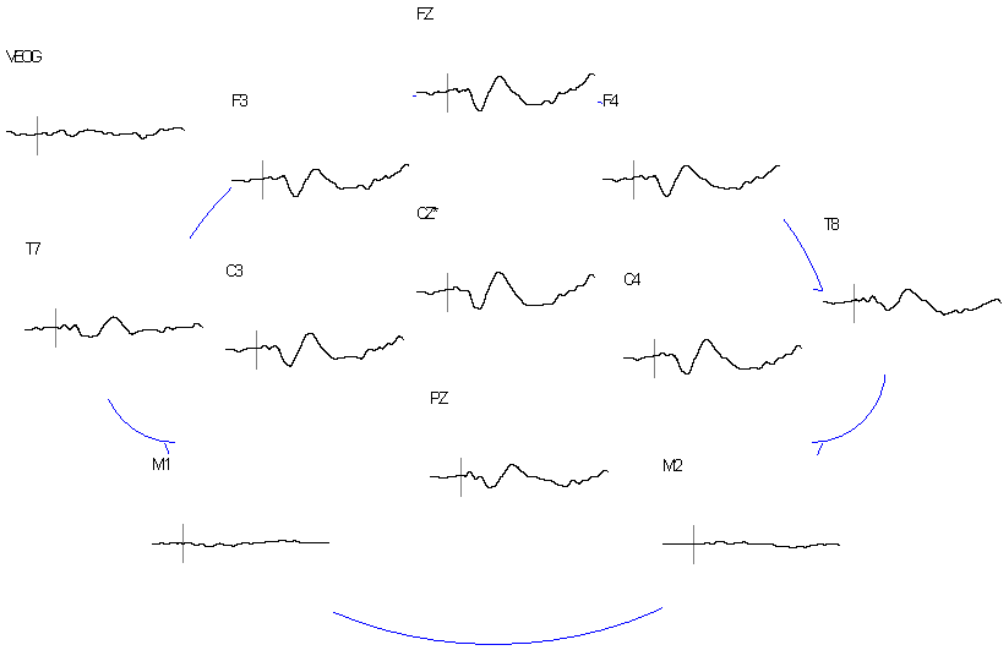


Figure 34. Grand average waveforms recorded across all electrode sites for male participants elicited via natural male /a/ tokens.

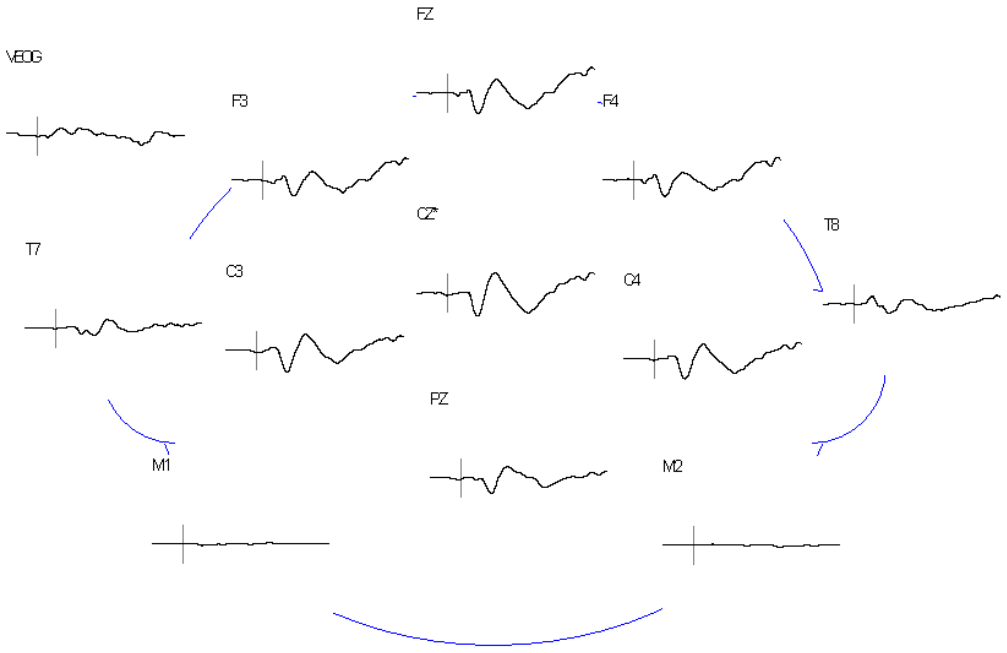


Figure 35. Grand average waveforms recorded across all electrode sites for female participants elicited via natural male /a/ tokens.

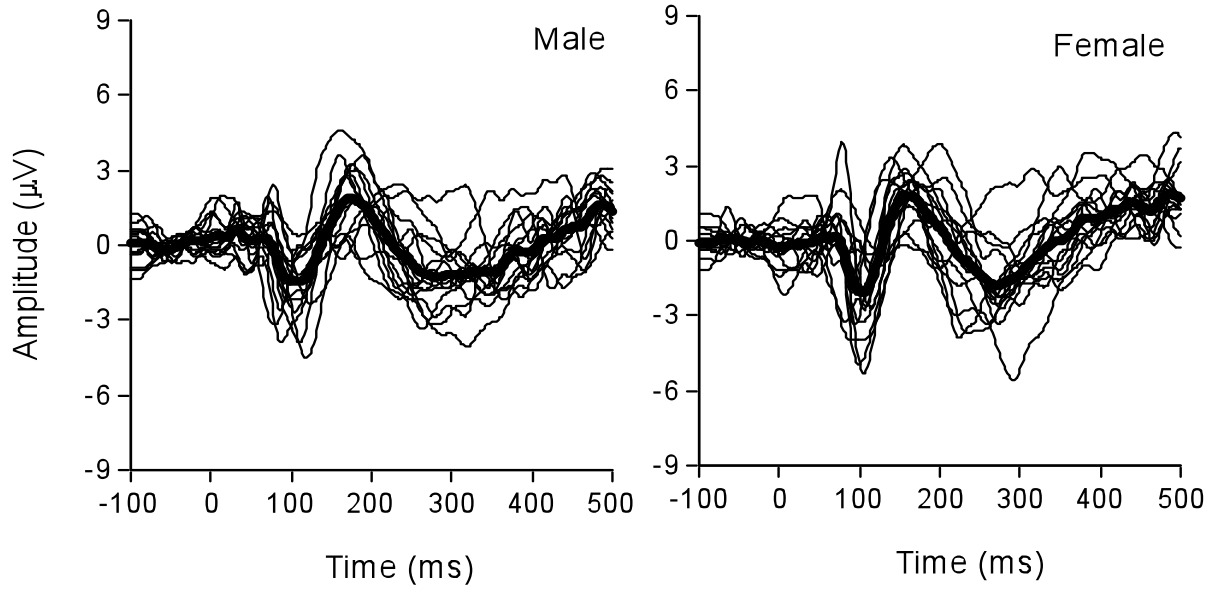


Figure 36. Individual participant waveforms recorded at Cz for male and female participants elicited via natural male /a/ tokens as a function of amplitude (μV) and time (ms). The grand average waveforms are displayed as the thickest black line.

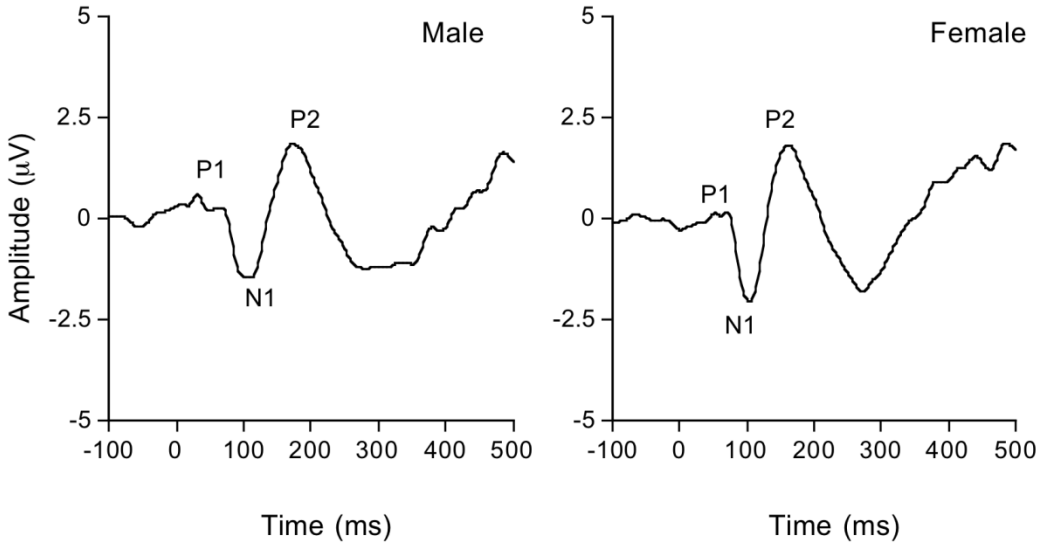


Figure 37. Grand average waveforms recorded at Cz for male and female participants elicited via natural male /a/ tokens as a function of amplitude (μV) and time (ms).

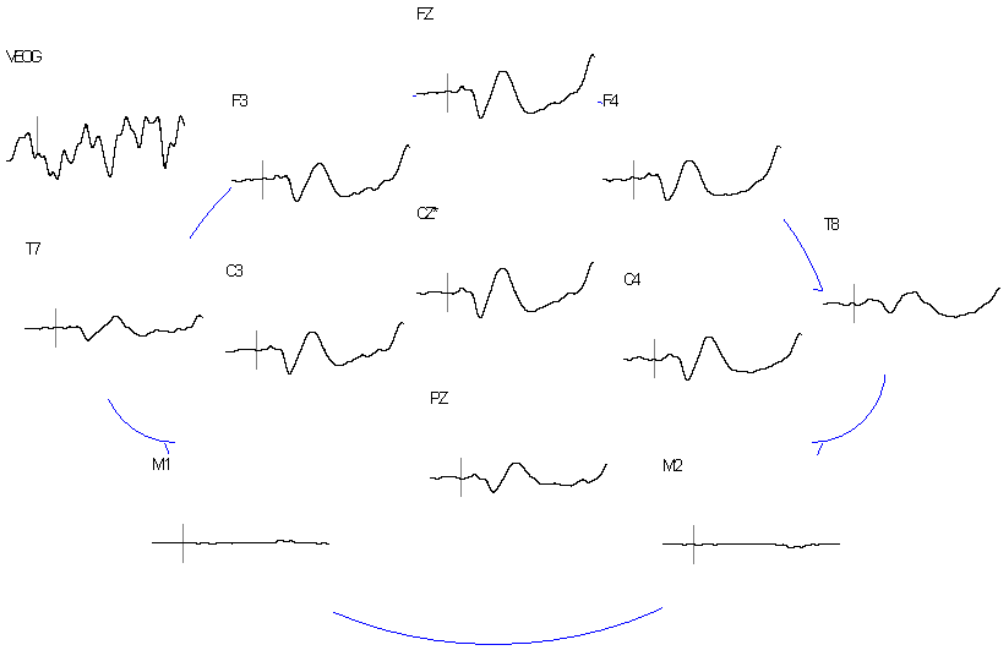


Figure 38. Grand average waveforms recorded across all electrode sites for male participants elicited via natural female /a/ tokens.

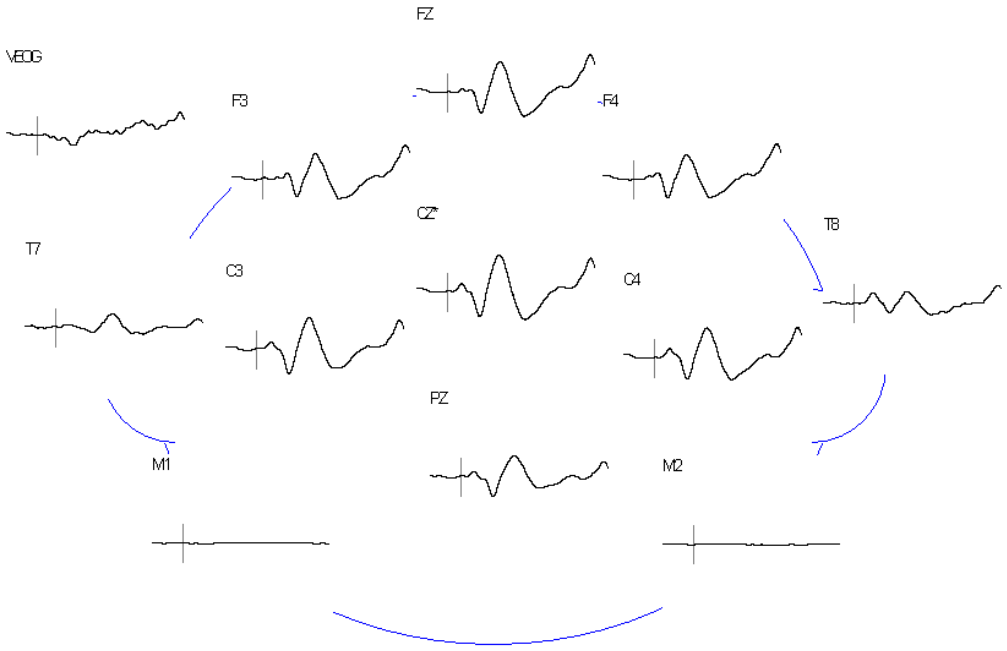


Figure 39. Grand average waveforms recorded across all electrode sites for female participants elicited via natural female /a/ tokens.

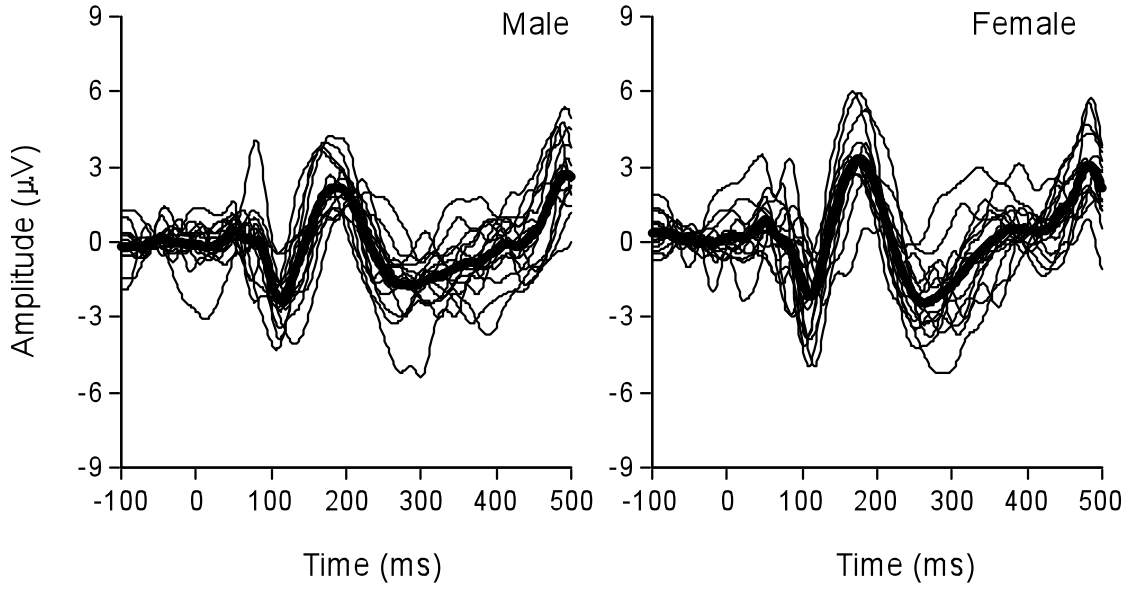


Figure 40. Individual participant waveforms recorded at Cz for male and female participants elicited via natural female /a/ tokens as a function of amplitude (μV) and time (ms). The grand average waveforms are displayed as the thickest black line.

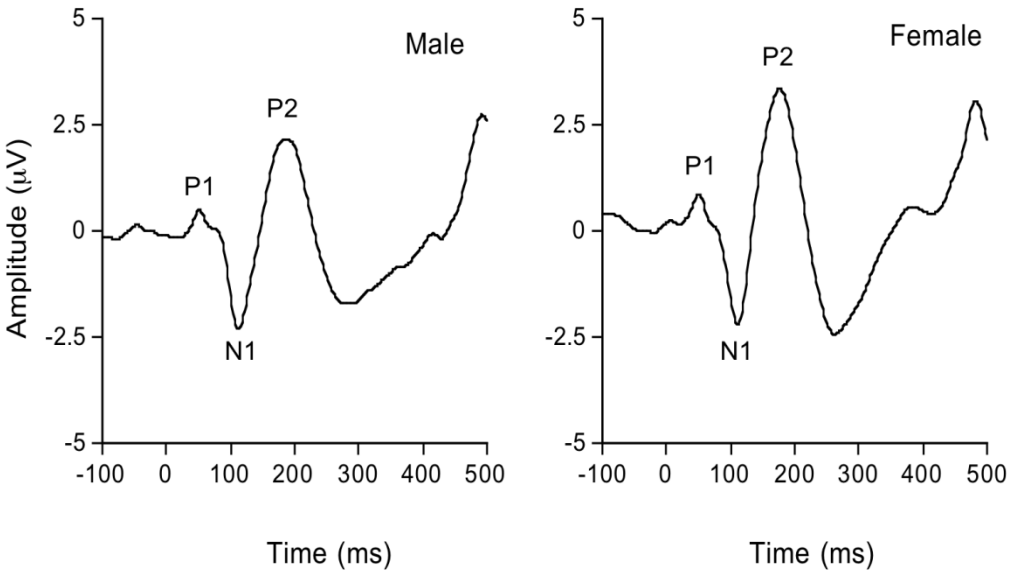


Figure 41. Grand average waveforms recorded at Cz for male and female participants elicited via natural female /a/ tokens as a function of amplitude (μV) and time (ms).

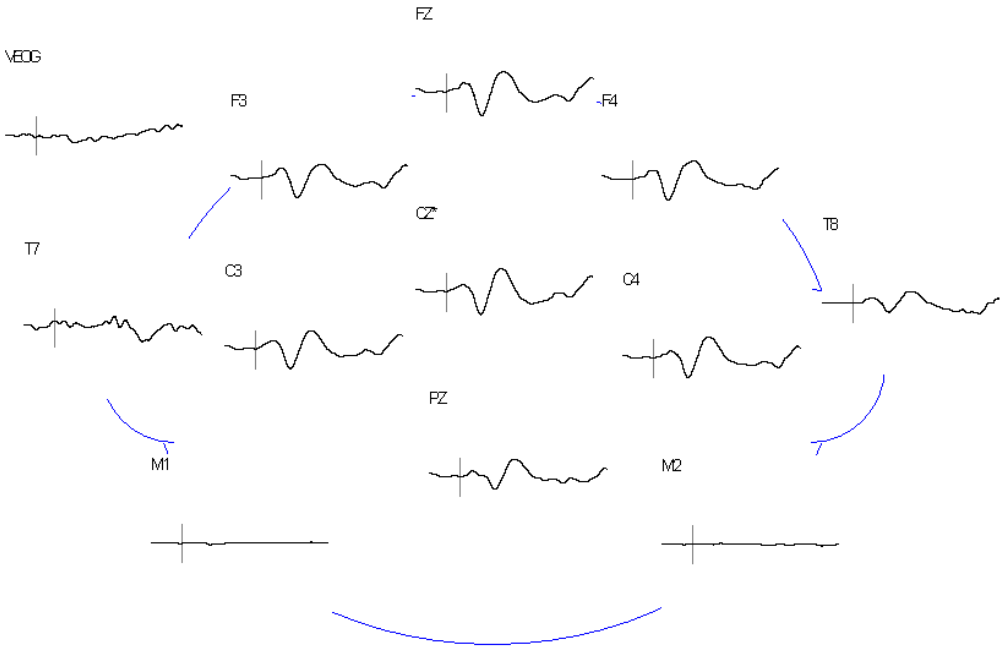


Figure 42. Grand average waveforms recorded across all electrode sites for male participants elicited via synthetic male /a/ tokens.

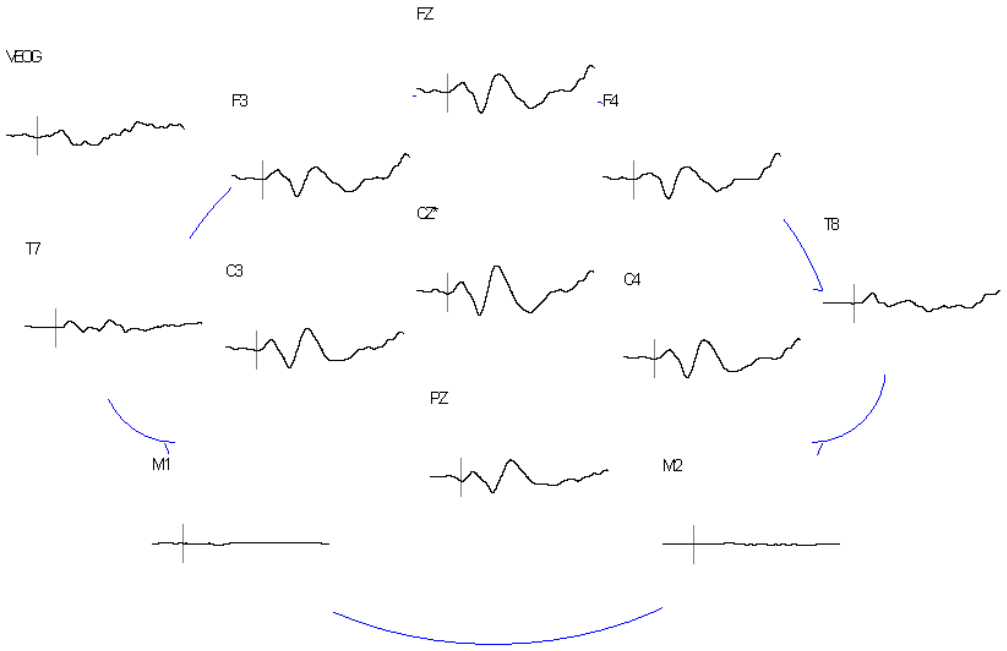


Figure 43. Grand average waveforms recorded across all electrode sites for female participants elicited via synthetic male /a/ tokens.

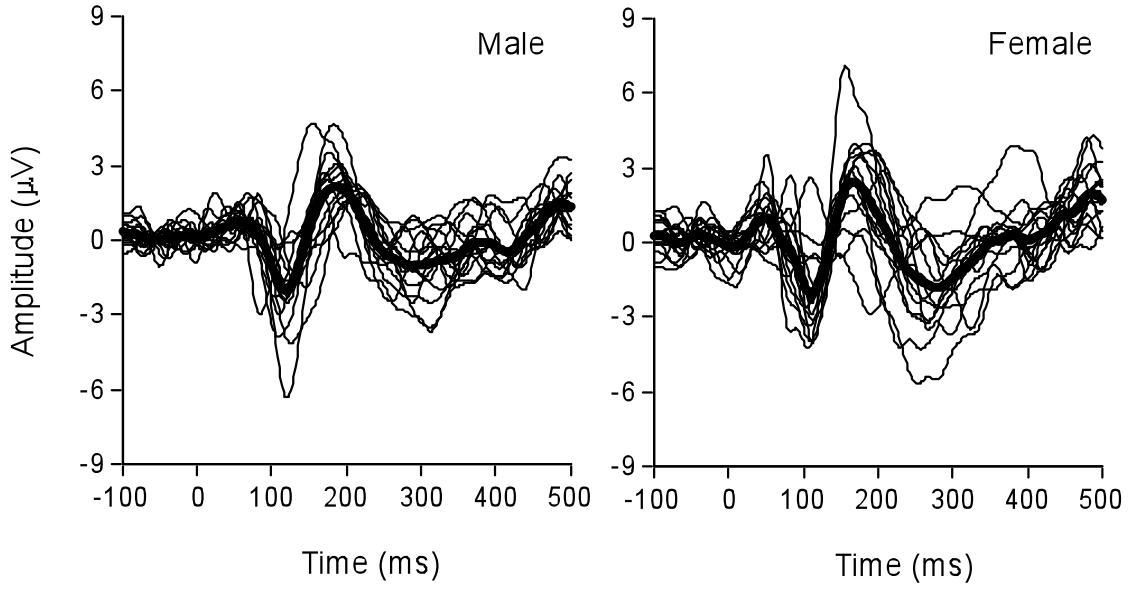


Figure 44. Individual participant waveforms recorded at Cz for male and female participants elicited via synthetic male /a/ tokens as a function of amplitude (μV) and time (ms). The grand average waveforms are displayed as the thickest black line.

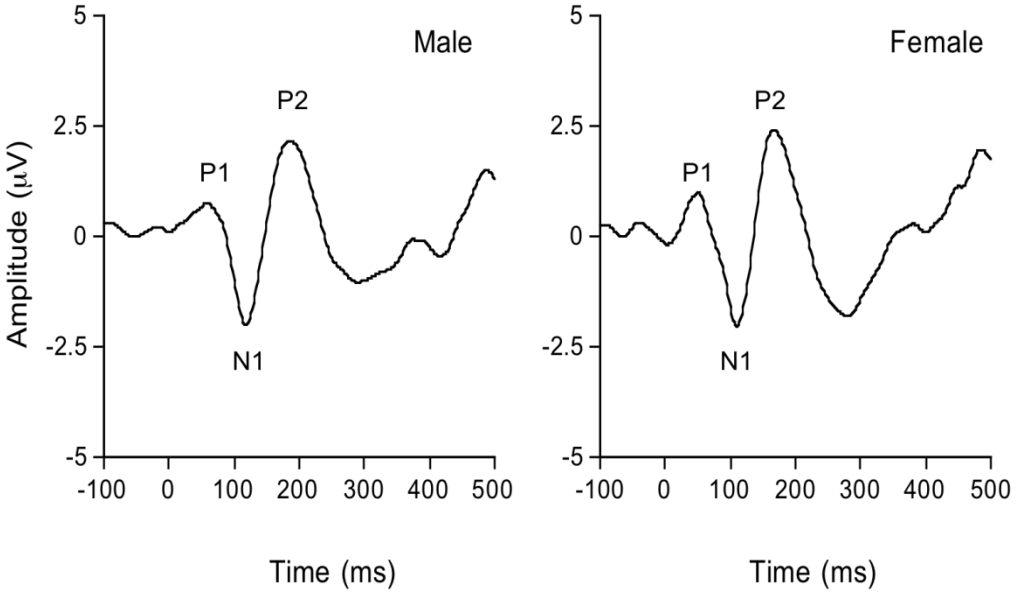


Figure 45. Grand average waveforms recorded at Cz for male and female participants elicited via synthetic male /a/ tokens as a function of amplitude (μV) and time (ms).

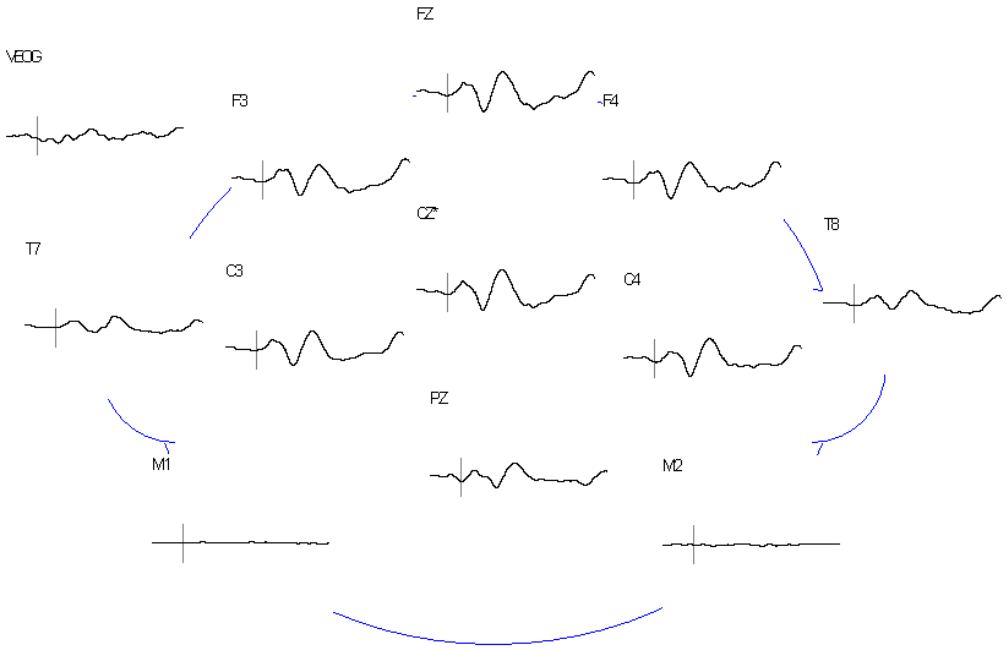


Figure 46. Grand average waveforms recorded across all electrode sites for male participants elicited via synthetic female /a/ tokens.

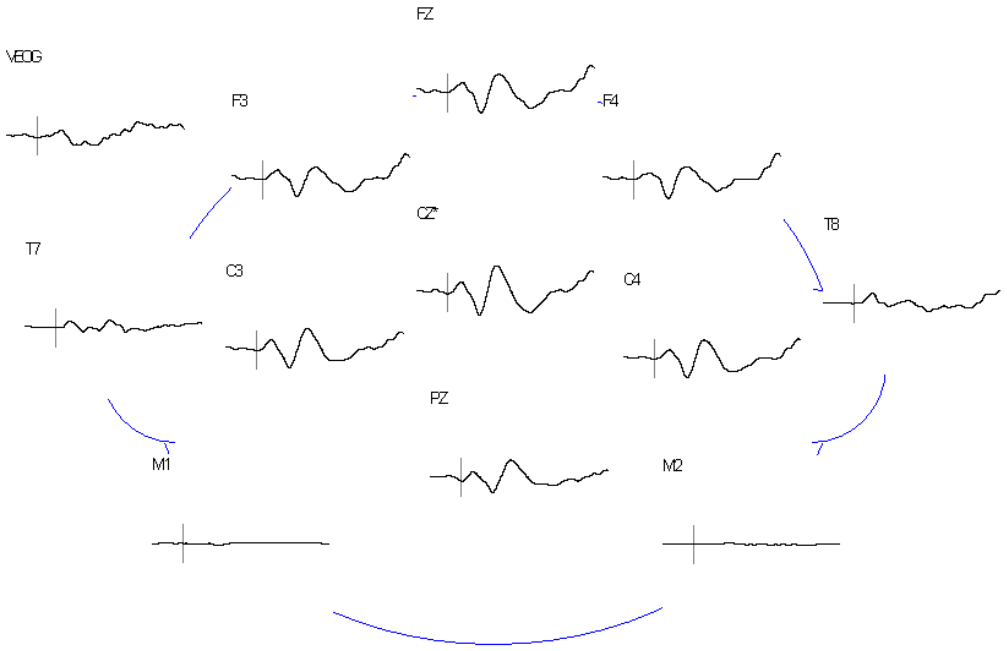


Figure 47. Grand average waveforms recorded across all electrode sites for female participants elicited via synthetic female /a/ tokens.

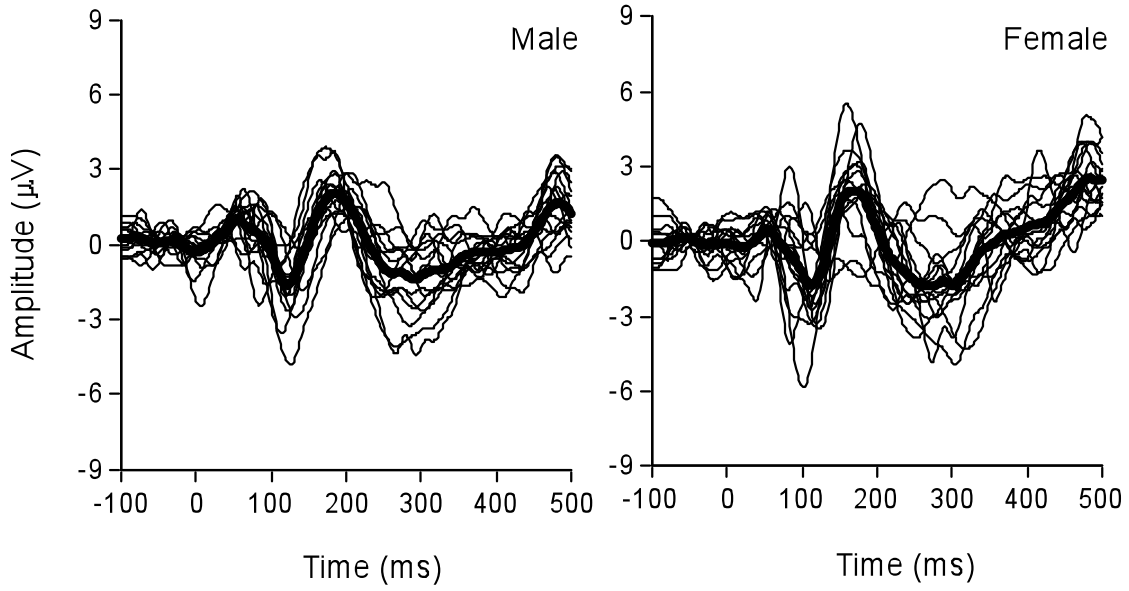


Figure 48. Individual participant waveforms recorded at Cz for male and female participants elicited via synthetic female /a/ tokens as a function of amplitude (μV) and time (ms). The grand average waveforms are displayed as the thickest black line.

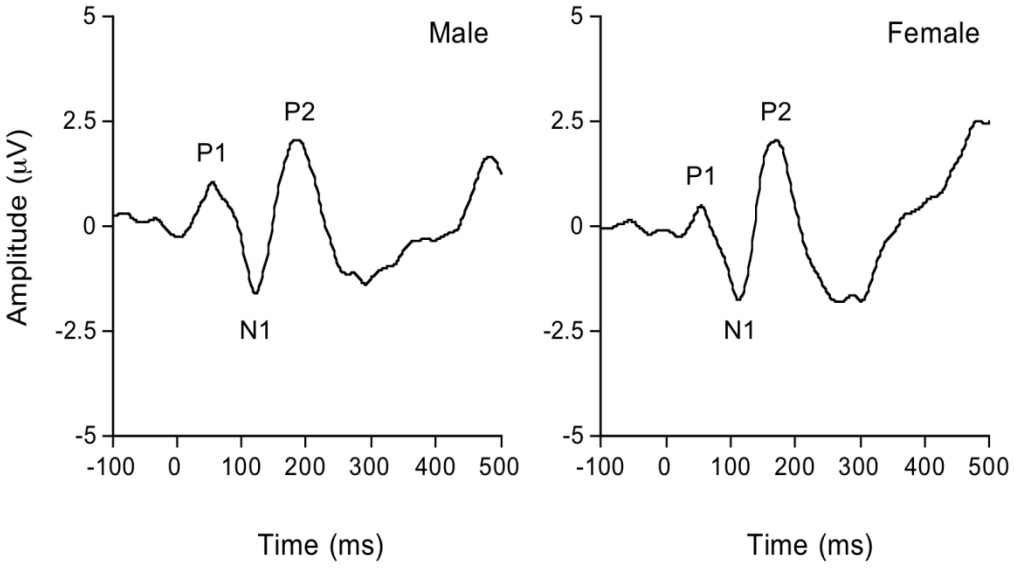


Figure 49. Grand average waveforms recorded at Cz for male and female participants elicited via synthetic female /a/ as a function of amplitude (μV) and time (ms).

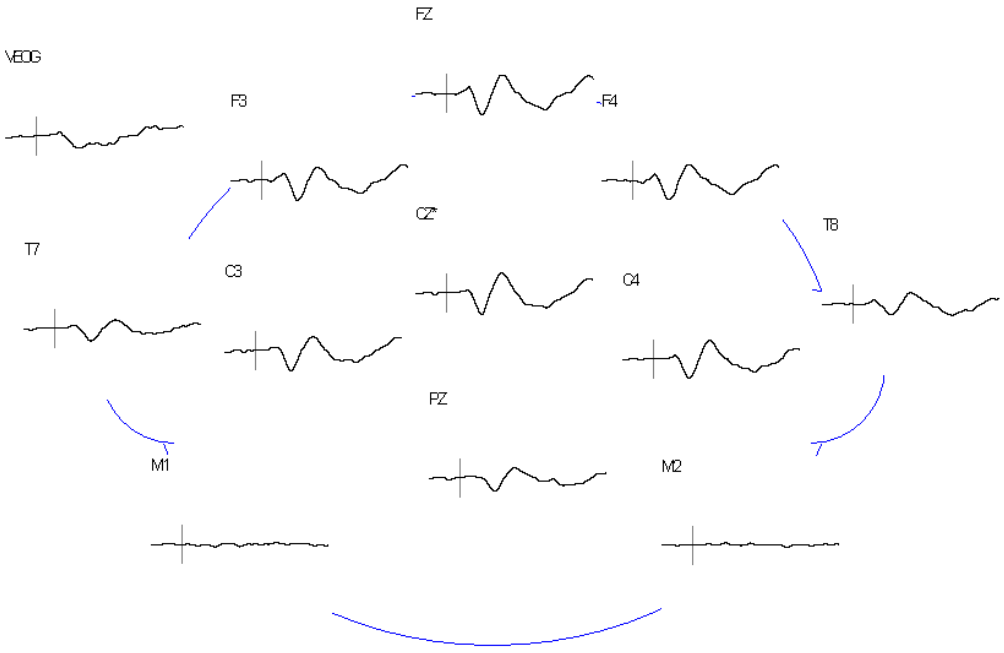


Figure 50. Grand average waveforms recorded across all electrode sites for male participants elicited via self-produced /a/ tokens.

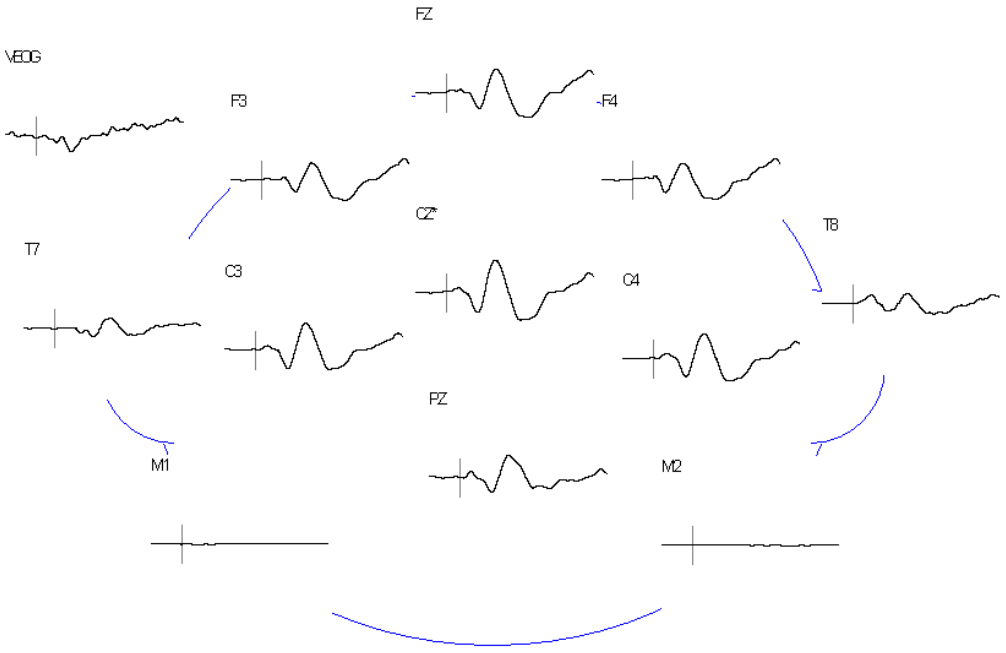


Figure 51. Grand average waveforms recorded across all electrode sites for female participants elicited via self-produced /a/ tokens.

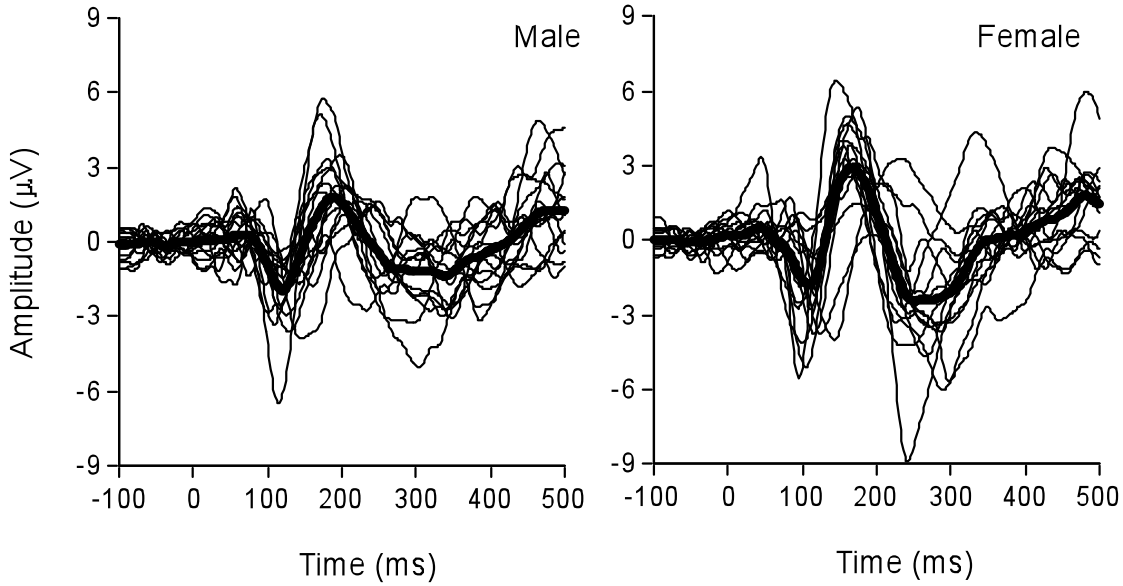


Figure 52. Individual participant waveforms recorded at Cz for male and female participants elicited via self-produced /a/ tokens as a function of amplitude (μV) and time (ms). The grand average waveforms are displayed as the thickest black line.

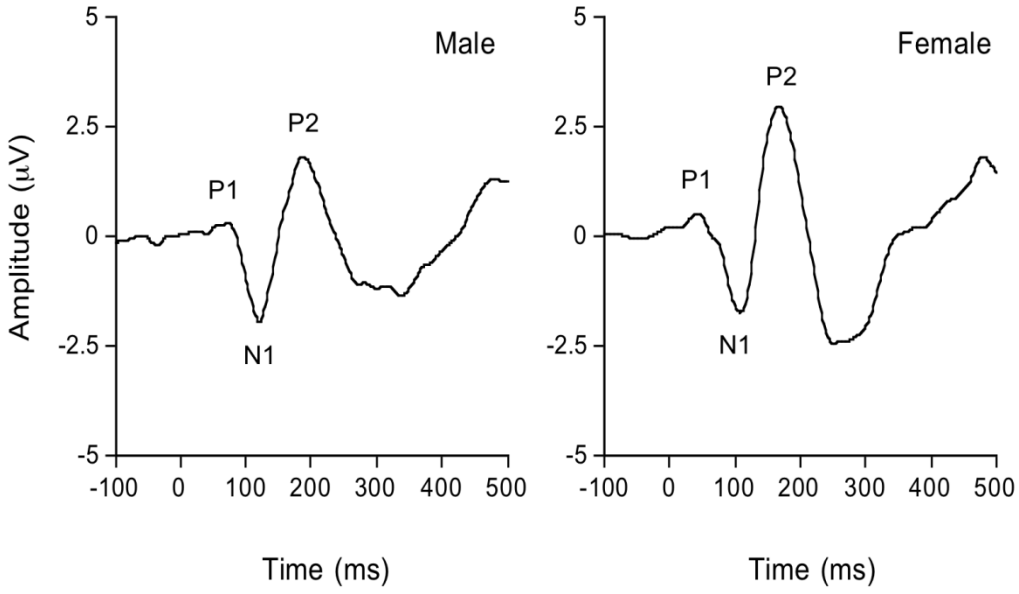


Figure 53. Grand average waveforms recorded at Cz for male and female participants elicited via self-produced /a/ tokens as a function of amplitude (μV) and time (ms).

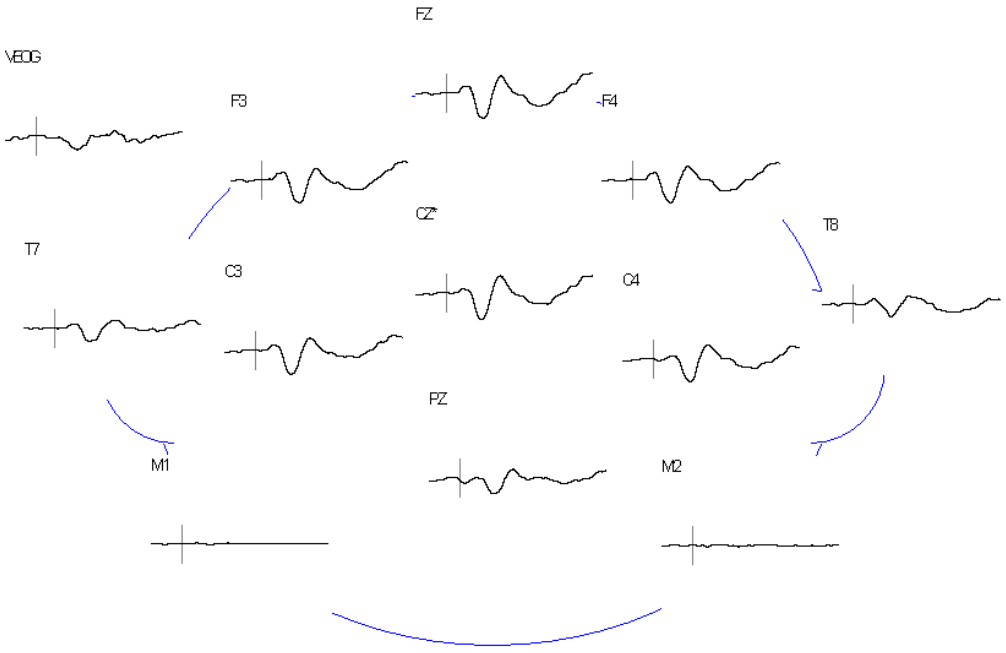


Figure 54. Grand average waveforms recorded across all electrode sites for male participants elicited via frequency shifted self-produced /a/ tokens.

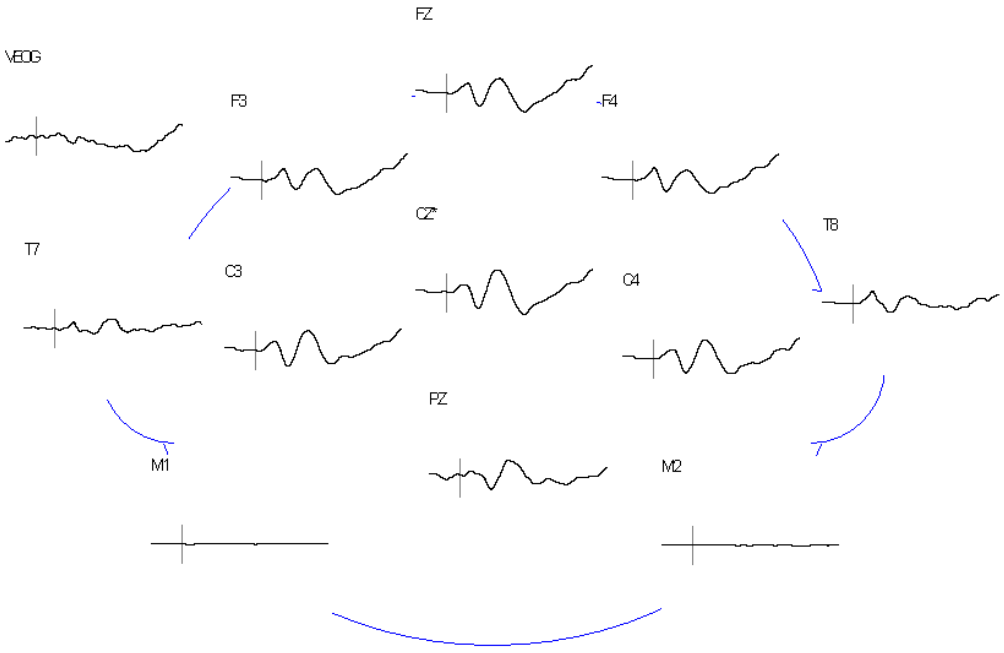


Figure 55. Grand average waveforms recorded across all electrode sites for female participants elicited via frequency shifted self-produced /a/ tokens.

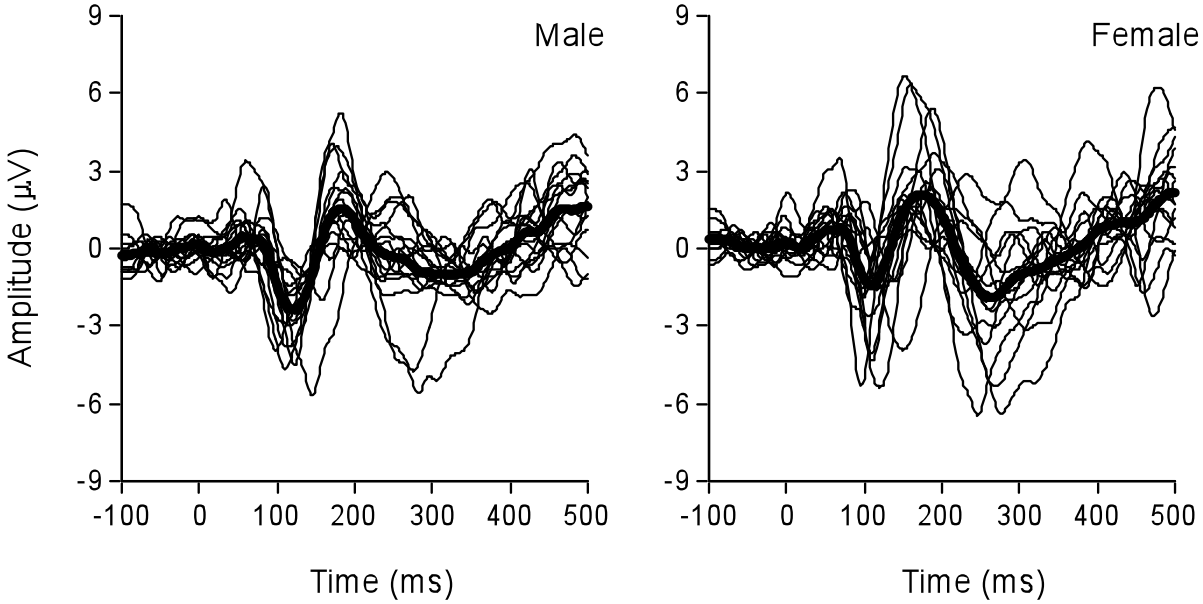


Figure 56. Individual participant waveforms recorded at Cz for male and female participants elicited via frequency shifted self-produced /a/ tokens as a function of amplitude (µV) and time (ms). The grand average waveforms are displayed as the thickest black line.

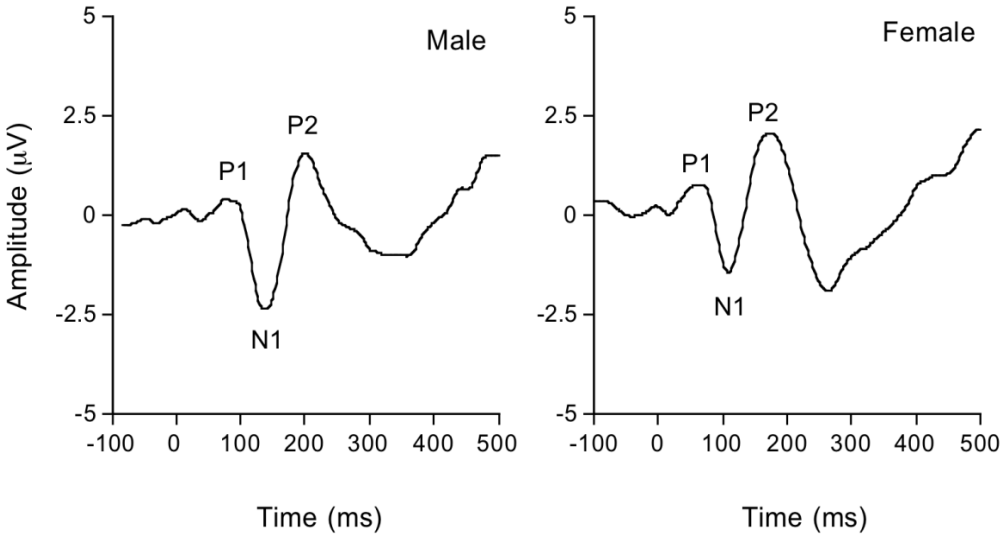


Figure 57. Grand average waveforms recorded at Cz for male and female participants elicited via frequency shifted self-produced /a/ tokens as a function of amplitude (μV) and time (ms).

contrasts are presented in Table 4. As presented in this table, significantly shorter P1 latencies occurred for the tonal stimulus versus the speech stimuli, for natural versus synthetic speech tokens, and for male tokens versus female tokens ($p < .05$).

N1 Latency

Mean latencies and standard deviations for the N1 waveform component, as a function of stimulus condition and gender, are presented in Table 5. Mean peak latencies and standard deviations for N1 as a function of stimulus condition presented are Figure 59. A two-factor mixed ANOVA was undertaken to examine the effect of stimulus condition and gender. The ANOVA summary for the N1 wave component is presented in Table 6. As evident in Table 6, only a main effect of condition was found for the N1 wave component.

To investigate the source of the main effect of stimulus condition on N1, five orthogonal single-*df* contrasts of interest were undertaken (Keppel & Wickens, 2004). Contrast summaries are presented in Table 7. As evident in this table significantly shorter N1 latencies occurred for the tonal stimulus versus the speech stimuli, natural versus synthetic speech tokens, and self-produced versus other speech ($p < .05$).

P2 Latency

Mean latencies and standard deviations for the P2 waveform component, as a function of condition and gender, are presented in Table 8 and Figure 60. A two-factor mixed ANOVA was undertaken to examine the effect of stimulus condition and gender. Significant effects of stimulus condition and gender were found (see Table 9). Females had significantly shorter latencies than males ($p = .007$).

To investigate the source of the main effect of stimulus condition on P2, five orthogonal single-*df* contrasts of interest were undertaken (Keppel & Wickens, 2004). A summary of those

Table 2.

Mean P1 Latencies and Standard Deviations as a Function of Stimulus Condition and Gender.

Stimuli	Latency (ms)	
	Male	Female
Tone	64.4	62.9
	(13.8)	(13.9)
Natural Male	64.1	64.9
	(13.0)	(12.5)
Natural Female	68.8	71.6
	(15.5)	(17.9)
Synthetic Male	72.3	75.2
	(12.7)	(27.6)
Synthetic Female	82.1	69.2
	(14.0)	(21.8)
Self-produced	77.2	69.5
	(13.6)	(18.4)
Frequency Shifted Self-produced	76.4	72.8
	(16.3)	(17.5)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

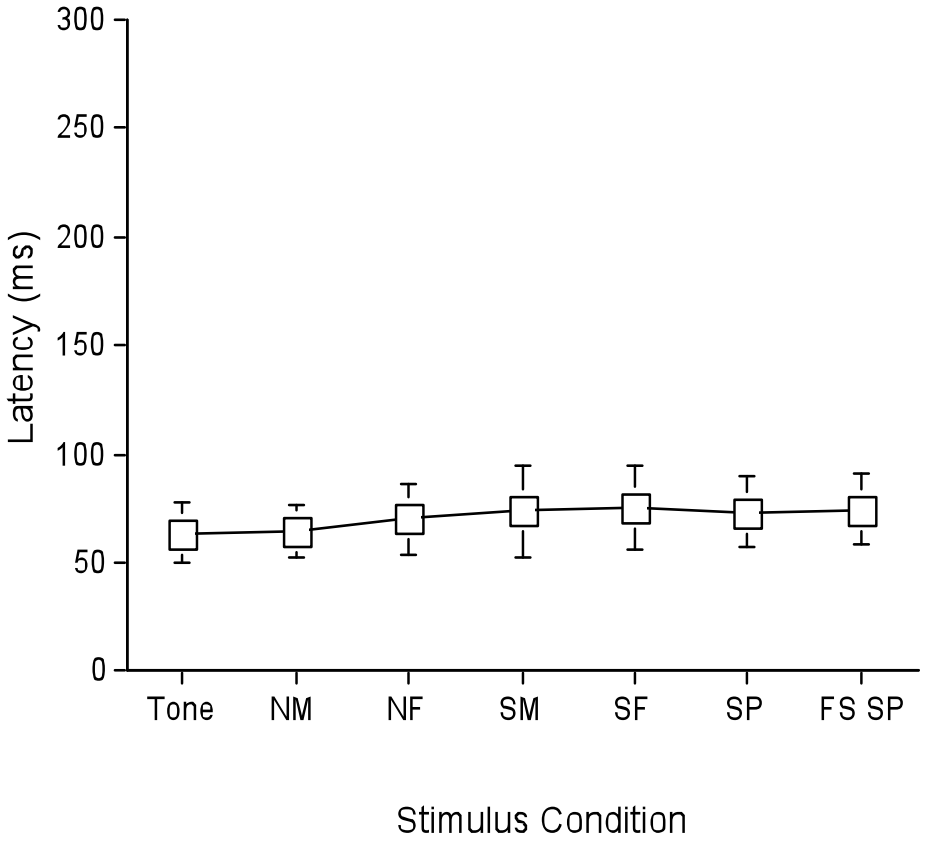


Figure 58. P1 mean peak latencies collapsed across gender a function of condition, Tone Burst, Natural Male (NM), Natural Female (NF), Synthetic Male (SM), Synthetic Female (SF), Self-produced (SP), and Frequency Shifted Self-produced (FS SP). Error bars represent +/- one *SD* of the mean.

Table 3.

Summary of Two-Factor Mixed ANOVA Comparing Differences Between P1 Mean Peak Latency (ms) as a Function of Condition (i.e., Tone Burst, Natural Male, Natural Female, Synthetic Male, Synthetic Female, Self-produced, and Frequency Shifted Self-produced) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	4309.79	6	1083.11	4.02	.0004* [a]	.13
Gender	394.97	1	394.97	.44	.51	.02
Gender X Condition	1549.56	6	389.45	1.44	.22[a]	.05

Note. * significant at $p < .05$; [a] Greenhouse-Geyser value.

Table 4.

Orthogonal Single-df Contrasts Investigating The Effect Of Condition on P1 Latency.

Contrast	*p*	η^2
Tone vs. Speech	<.0001*	.35
Synthetic vs. Natural	.029*	.15
Male tokens vs. Female tokens	.042*	.14
Self-produced vs. Other Speech	.24	.006
Self-produced vs. Frequency Shifted Self-produced	.69	.56

Note. * significant at $p < .05$.

Table 5.

Mean N1 Peak Latencies and Standard Deviations as a Function of Condition and Gender.

Stimuli	Latency (ms)	
	Male	Female
Tone	103.1	98.4
	(12.0)	(7.5)
Natural Male	103.1	102.23
	(13.0)	(13.7)
Natural Female	113.6	109.89
	(9.7)	(11.3)
Synthetic Male	115.7	119.5
	(14.4)	(22.7)
Synthetic Female	119.9	110.9
	(14.2)	(17.6)
Self-produced	123.7	109.7
	(16.3)	(13.9)
Frequency Shifted Self-produced	122.1	113.1
	(15.5)	(17.2)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

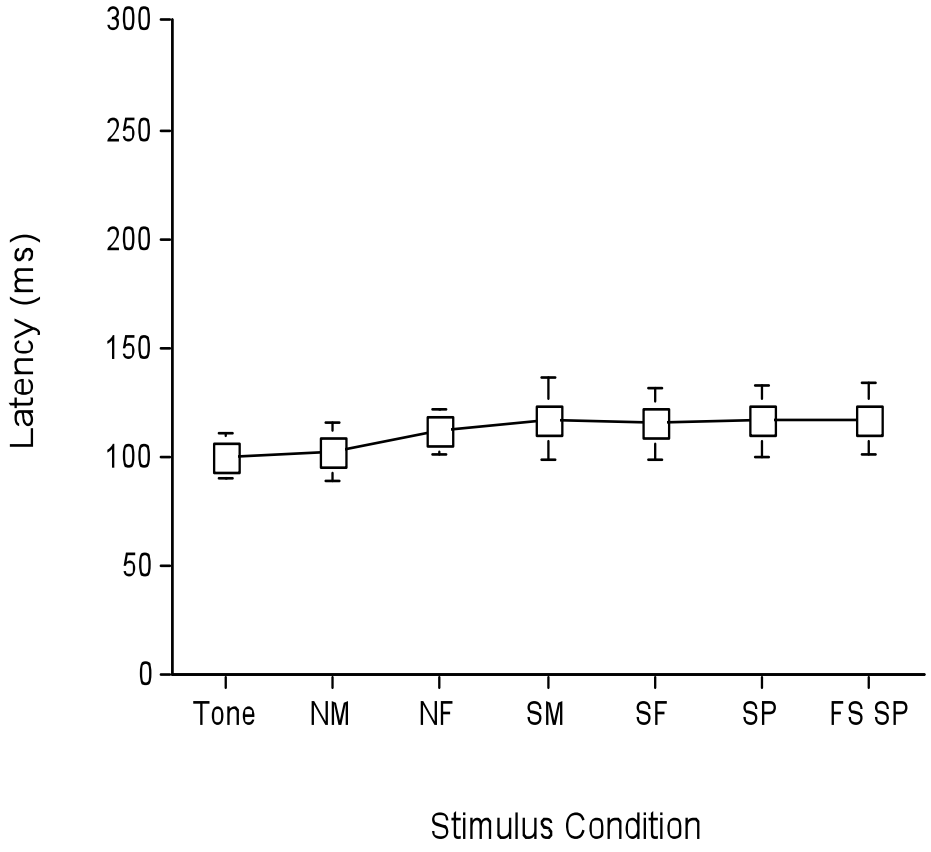


Figure 59. N1 mean peak latencies collapsed across gender as a function of condition, Tone Burst, Natural Male (NM), Natural Female (NF), Synthetic Male (SM), Synthetic Female (SF), Self-produced (SP), and Frequency Shifted Self-produced (FS SP). Error bars represent +/- one *SD* of the mean.

Table 6.

Summary of Two-Factor Mixed ANOVA Comparing Differences Between N1 Mean Peak Latency (ms) as a Function of Condition (i.e., Tone Burst, Natural Male, Natural Female, Synthetic Male, Synthetic Female, Self-produced, and Frequency Shifted Self-produced) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	9314.05	6	2340.44	12.54	<.0001* [a]	.31
Gender	1504.02	1	1504.02	1.98	.17	.07
Gender X Condition	1558.25	6	391.56	2.10	.086[a]	.07

Note. * significant at $p < .05$; [a] Greenhouse-Geisser value.

Table 7.

Orthogonal Single-df Contrasts Investigating The Effect Of Condition on N1 Latency.

Contrast	*p*	η^2
Tone vs. Speech	<.0001*	.62
Synthetic vs. Natural	<.0001*	.40
Male tokens vs. Female tokens	.12	.083
Self-produced vs. Other Speech	.043*	.133
Self-produced vs. Frequency Shifted Self-produced	.68	.43

Note. * significant at $p < .05$.

contrasts as related to P2 component latency is presented in Table 10. As evident in this table, significantly shorter P2 latencies occurred for the tonal stimulus versus the speech stimuli and for natural versus synthetic speech tokens ($p < .0001$).

P1-N1 Amplitude

Mean peak amplitudes and standard deviations, as a function of stimulus condition and gender, for the P1-N1 waveform complex are presented in Table 11. A separate two-factor mixed ANOVA was undertaken to examine the effect of stimulus condition and gender on P1-N1 waveform amplitude. The ANOVA summary for the P1-N1 wave component analyses are presented in Tables 12. As evident in Table 12, there were no significant main effects for P1-N1 amplitude. The P1-N1 amplitude grand mean was 3.49 μV (95% CI 3.01-3.96).

N1-P2 Amplitude

Mean peak amplitudes and standard deviations, as a function of stimulus condition and gender, for the N1-P2 waveform complex are presented in Table 13. Mean peak amplitudes and standard deviations for the N1-P2 complex are presented in Figure 61.

A separate two-factor mixed ANOVA was undertaken to examine the effect of stimulus condition and gender on N1-P2 waveform amplitude. The ANOVA summary for N1-P2 wave component analyses is presented in Tables 14. Only a significant main effect of stimulus condition was found (see Table 14). To investigate the source of the main effect of stimulus condition on N1-P2 amplitude, five orthogonal single-*df* contrasts, as employed above, were undertaken. Summaries of those contrasts are presented in Table 15. The only significant difference was significantly larger N1-P2 amplitudes for the tonal stimulus versus the speech stimuli ($p < .0001$).

Table 8.

Mean P2 Peak Latencies and Standard Deviations as a Function of Condition and Gender.

Stimuli	Latency (ms)	
	Male	Female
Tone	164.8	158.7
	(11.2)	(9.1)
Natural Male	176.1	164.0
	(26.3)	(18.9)
Natural Female	186.8	178.9
	(17.5)	(12.0)
Synthetic Male	190.5	176.8
	(19.8)	(25.2)
Synthetic Female	187.2	172.8
	(12.9)	(16.5)
Self-produced	194.9	170.8
	(28.7)	(20.1)
Frequency Shifted Self-produced	188.4	172.1
	(24.5)	(23.3)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

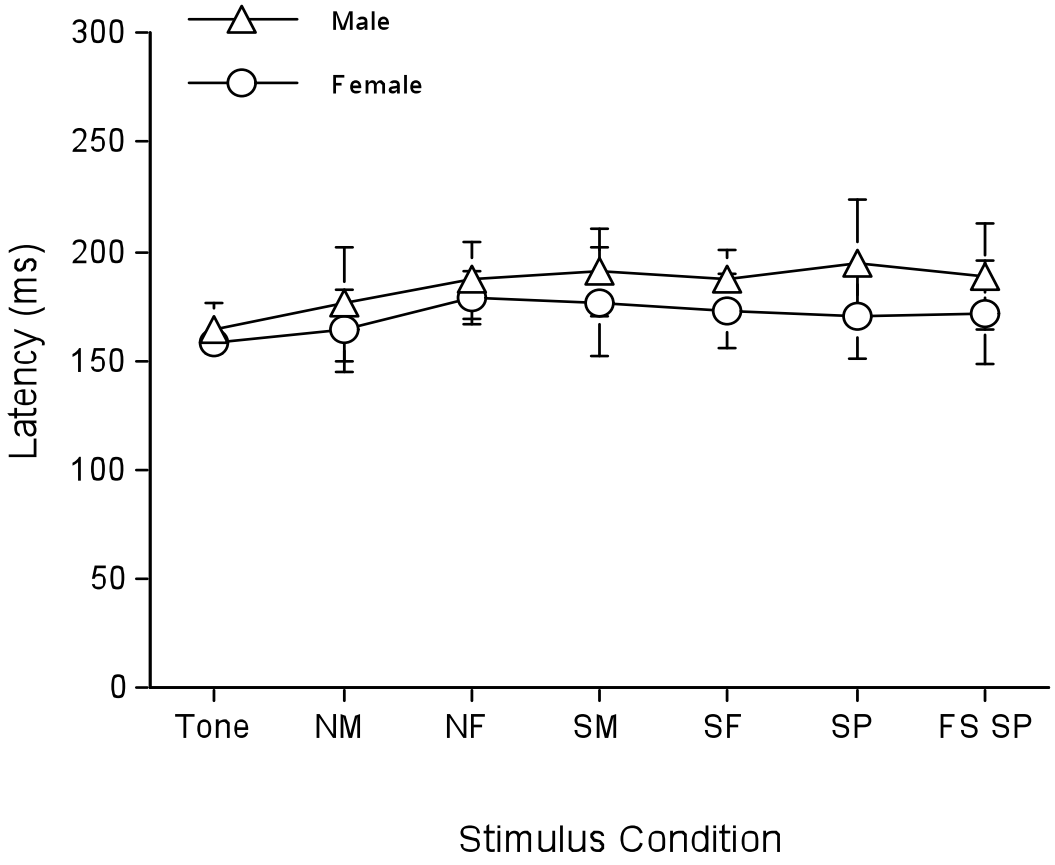


Figure 60. P2 mean peak latencies as a function of gender and stimulus condition, Tone Burst, Natural Male (NM), Natural Female (NF), Synthetic Male (SM), Synthetic Female (SF), Self-produced (SP), and Frequency Shifted Self-produced (FS SP). Error bars represent +/- one *SD* of the mean.

Table 9.

Summary of Two-Factor Mixed ANOVA Comparing Differences Between P2 Mean Peak Latency (ms) as a Function of Condition (i.e., Tone Burst, Natural Male, Natural Female, Synthetic Male, Synthetic Female, Self-produced, and Frequency Shifted Self-produced) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	12396.72	6	3078.52	7.62	<.0001* [a]	.21
Gender	9601.91	1	9601.91	8.42	.007*	.23
Gender X Condition	1570.90	6	390.15	.97	.43[a]	.03

Note. * significant at $p < .05$; [a] Greenhouse-Geisser value.

Table 10.

Orthogonal Single-df Contrasts Investigating The Effect Of Stimulus Condition on P2 Latency.

Contrast	*p*	η^2
Tone vs. Speech	<.0001*	.62
Synthetic vs. Natural	.042*	.14
Male tokens vs. Female tokens	.075	.11
Self-produced vs. Other Speech	.46	.019
Self-produced vs. Frequency Shifted Self-produced	.51	.015

Note. * significant at $p < .05$.

Table 11.

Mean P1-N1 Amplitudes and Standard Deviations for as a Function of Condition and Gender.

Stimuli	Amplitude (μV)	
	Male	Female
Tone	3.4	3.9
	(1.5)	(1.9)
Natural Male	3.0	3.4
	(1.1)	(1.6)
Natural Female	3.4	3.7
	(1.1)	(1.7)
Synthetic Male	3.5	3.9
	(1.6)	(1.9)
Synthetic Female	2.9	3.3
	(1.6)	(1.6)
Self-produced	3.1	3.8
	(1.7)	(2.1)
Frequency Shifted Self-produced	3.6	3.8
	(1.6)	(2.0)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

Table 12.

Summary of Two-Factor Mixed ANOVA Comparing Differences Between P1-N1 Mean Amplitudes (μV) as a Function of Condition (i.e., Tone Burst, Natural Male, Natural Female, Synthetic Male, Synthetic Female, Self-produced, and Frequency Shifted Self-produced) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	10.78	6	1.80	1.31	.25	.045
Gender	8.03	1	8.03	.72	.40	.025
Gender X Condition	.92	6	.15	.11	1.00	.004

Note. * significant at $p < .05$.

Table 13.

Mean N1-P2 Amplitudes and Standard Deviations as a Function of Condition and Gender.

	Amplitude (μV)	
Stimuli	Male	Female
Tone	6.6	8.5
	(1.6)	(3.2)
Natural Male	4.4	4.8
	(2.2)	(2.0)
Natural Female	5.0	6.2
	(1.6)	(2.4)
Synthetic Male	5.2	5.3
	(2.1)	(2.6)
Synthetic Female	4.3	4.9
	(1.3)	(2.7)
Self-produced	5.0	6.3
	(2.5)	(2.6)
Frequency Shifted Self-produced	5.2	5.6
	(2.4)	(3.0)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

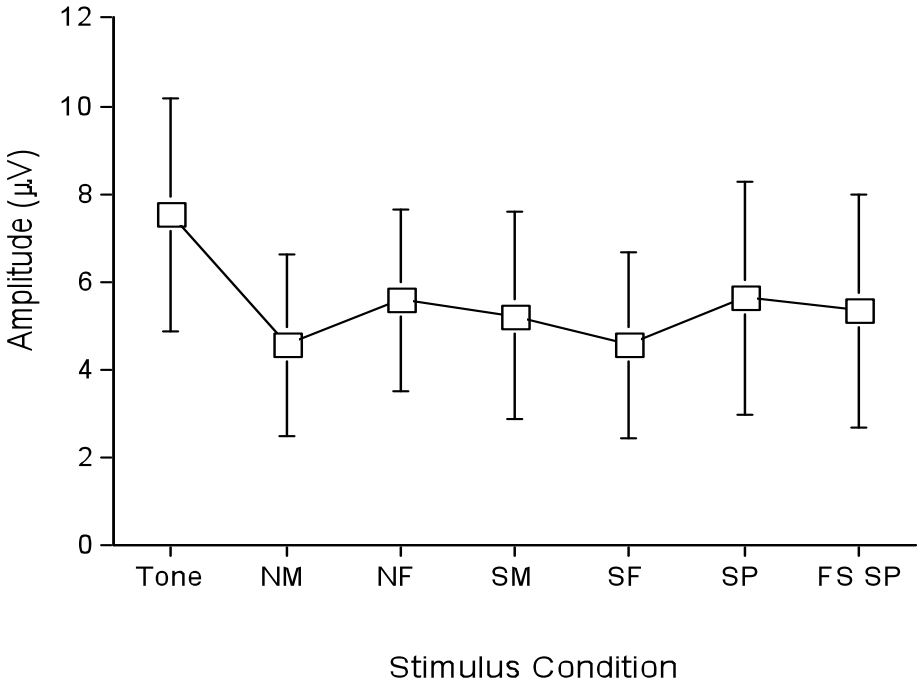


Figure 61. N1-P2 complex mean amplitudes collapsed across gender as a function and stimulus condition, Tone Burst, Natural Male (NM), Natural Female (NF), Synthetic Male (SM), Synthetic Female (SF), Self-produced (SP), and Frequency Shifted Self-produced (FS SP). Error bars represent +/- one *SD* of the mean.

Table 14.

Summary of Two-Factor Mixed ANOVA Comparing Differences Between N1-P2 Mean Amplitudes (μV) as a Function of Condition (i.e., Tone Burst, Natural Male, Natural Female, Synthetic Male, Synthetic Female, Self-produced, and Frequency Shifted Self-produced) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	179.63	6	29.94	12.68	<.0001*	.31
Gender	36.76	1	36.76	1.48	.23	.05
Gender X Condition	19.68	6	3.28	1.39	.22	.047

Note. * significant at $p < .05$.

Table 15.

Orthogonal Single-df Contrasts Investigating The Effect Of Stimulus Condition on N1-P2 Amplitude.

Contrast	*p*	η^2
Tone vs. Speech	<.0001*	.59
Synthetic vs. Natural	.47	.018
Male tokens vs. Female tokens	.53	.014
Self-produced vs. Other Speech	.11	.086
Self-produced vs. Frequency Shifted Self-produced	.42	.023

Note. * significant at $p < .05$.

Discussion

Effect of Stimulus Condition

To better understand auditory monitoring of exogenously presented stimuli during passive listening and to investigate central auditory processing via ERPs, the following general experimental was question asked: Are auditory P1-N1-P2 component latencies and amplitudes affected by stimulus condition? To address this question, the effect of stimulus condition on passive auditory cortical processing as demonstrated by P1-N1-P2 component latencies and amplitudes were evaluated using both non-speech and speech stimuli. In general, it was hypothesized that responses evoked with the non-speech token will differ from those recorded with the speech tokens. Specifically, component latencies will be shorter and amplitudes larger when evoked with tones compared to speech (Cëponiene et al., 2001; Tiitinen et al., 1999). Additionally, it was predicted that no significant differences will be seen across ERPs elicited via speech tokens.

Nonspeech versus speech tokens. Consistent with the proposed experimental predictions P1-N1-P2 component latencies were significantly shorter when evoked via the tonal stimulus compared to the speech stimulus. Notably, these findings were consistent with the findings of other investigators. Cëponiene et al. (2001) and Tiitinen et al. (1999) reported shorter latency values for waveform components elicited with tonal stimuli versus speech stimuli. In terms of P1-N1 and N1-P2 amplitudes, no significant differences were found for the P1-N1complex; however, significantly larger N1-P2 amplitudes were recorded with the tone burst than the speech tokens. Again, this finding is consistent with the findings of others (Tiitinen et al., 1999).

Proponents of the "speech is special" philosophy could point to the fact that cortical processing of tonal stimuli differs from that of speech tokens as evidence for their theory

(Liberman & Mattingly, 1985, 1989). However, since acoustical differences in the temporal and spectral domains were evident one cannot attribute component differences to the dichotomy of speech versus non-speech processing.

It is established that stimulus characteristics such as duration, rise time, intensity, and frequency affect P1-N1-P2 responses. In looking at the experimental stimuli with respect to acoustical parameters the most notable difference between the non-speech and speech tokens was the difference in token duration. The duration of the tone burst was 70 ms time while the duration of all speech tokens were held constant at approximately 400 ms. Investigators have shown differences in component latencies as a function of duration and rise/fall time (Davis & Zerlin, 1966; Eddins & Peterson, 1999; Eulitz et al., 1995; Onishi & Davis, 1968; Skinner & Jones, 1968). For example, Onishi and Davis (1968) and Skinner and Jones (1968) reported that as the duration of tonal stimulus decreased N1 and P2 latency increased, while component amplitude increased as duration decreased. The duration of the 70 ms tone burst is obviously shorter than that of the speech tokens (i.e., 400 ms); however, the P1-N1-P2 component latencies were significantly shorter when elicited via the tone burst compared to the speech tokens. It should be noted; however, that Onishi and Davis found the effect of stimulus duration on N1-P2 latencies were most significant at shorter durations. In that, as the duration of the tone increased beyond approximately 30 ms, stimulus duration ceased to have a significant effect on component latencies. In this experiment, therefore, the effect of stimulus duration should be minimal given that all stimuli were longer than 30 ms. That is differences in stimulus durations are not likely to contribute to the latency differences found in this experiment. Differences in token rise time on the other hand do affect component latency and amplitude such that these variations would reasonably account for the presented latency and amplitude differences.

With respect to stimulus onset the P1-N1-P2 complex, specifically the N1 component, is an onset response and therefore greatly impacted by the "onset" characteristics of the token (i.e., the initial 100 ms; Elfner, Gustafson, & Williams, 1976; Skinner & Jones, 1968). Thus, manipulating the onset of the stimulus (i.e., the rise time) can affect component latency and amplitude. B. A. Milner (1969) found as stimulus rise time increased component amplitude decreased. Onishi and Davis (1968) found that rise times shorter than 50 ms did not significantly affect N1-P2 amplitude; however as rise time was increased beyond 50 ms N1-P2 amplitude decreased. Thomson, Goswami, and Baldeweg (2009) also reported that N1 component amplitudes decreased as stimuli rise time increased. For component latency, Hyde (1997) reported that as stimulus rise time increased component latency increased. In this experiment, the rise time for the non-speech token was 10 ms (see Figure 2). The rise time of the speech stimuli was not controlled across tokens so the parameters of these tokens would remain as close to naturally occurring as possible. Therefore, rise time for the speech tokens varied. This was especially true for the participant self-produced tokens where the participant's self-produced vowel was presented as it was recorded. Rise times for the natural male and natural female tokens were approximately 53 ms and 37 ms, respectively (see Figures 3 and 4), 53 ms for the synthetic male token (see Figure 5), 27 ms for the synthetic female token (see Figure 6), 68 ms for the male self-produced example token (see Figure 7), 53 ms for the male frequency shifted self-produced example token (see Figure 8), 54 ms for the female self-produced example token (see Figure 9), and lastly 54 ms for the female frequency shifted self-produced example token (see Figure 10). It is possible that P1-N1-P2 latency and amplitude differences found here were wholly related to stimuli rise time differences, given that the onset of the non-speech tone was

comparative shorter and more rapid than the onset of the speech tokens, which were relative long.

Stimulus intensity is another acoustical parameter that affects component latencies. In this experiment all stimuli were calibrated and normalized to 75 dB pSPL. However, when considering the differences in acoustical parameters between the stimuli it is possible that the intensity of the stimuli were perceptually different as a function of temporal integration. In numerous studies shorter N1-P2 latencies have been reported as stimulus intensity increased (Antinoro et al., 1969; Beagley & Knight, 1967; Davis et al., 1966; Davis & Zerlin, 1966; Gerin et al., 1972; Onishi & Davis, 1968; Picton et al., 1977; Rapin et al., 1966; Rothman et al., 1970). Therefore, if the tone burst was perceptually louder than that of the speech stimuli, it stands to reason that the P1-N1-P2 latency differences may be attributed to latency/intensity effects.

With consideration to two key factors; however, it is believed that intensity effects were not solely responsible for the findings of this experiment. First, researchers have shown that changes in stimulus durations affect perceptual loudness, in that, as the duration of the stimulus decreases the overall perceptual intensity decreases as well (Moore, 1997). This phenomenon holds true for shorter duration stimuli (i.e., < 200 ms). In other words, in order for a long duration stimulus and short stimulus to be perceived as equally loud, the shorter duration stimuli must actually be more intense (Florentine, Fastl, & Buus, 1988; Garner & Miller, 1947; Plomp & Bouman, 1959). In terms of electrophysiological measures, again, researchers have shown a decrease in component latencies with increasing duration and intensity (Eddins & Peterson, 1999; Onishi & Davis, 1968).

Findings supporting an effect of temporal integration were not observed in the current experiment. Namely, temporal integration influences would be demonstrated as shorter P1-N1-

P2 latencies for the speech stimuli considering these tokens are longer in terms of duration than the tone bursts and in turn would be theatrically perceived as louder. In the current experiment the opposite held true. P1-N1-P2 component latencies were all significantly earlier when elicited via the 70 ms tone burst than those component latencies elicited with speech tokens.

As demonstrated by Eddins and Peterson (1999), the influence of temporal integration on component latency is more predominate at intensity levels close to listener threshold. In the current experiment, all tokens were calibrated to 75 dB pSPL, which is well above the threshold for normal hearing listeners. Given this, it is assumed that the influence of intensity and/or intensity with relation to stimulus duration is minimal to the current experimental findings.

Although, the frequency of the tone was constructed based on the center frequency of the synthetic vowel tokens, the difference between the responses may be related to the spectral characteristics of the tokens and the complexity of the tokens (Uppenkamp, Johnsrude, Norris, Marslen-Wilson, & Patterson, 2006). Additionally, ISI has been shown to affect ERP components. Given that the ISI was kept constant at 1.1/s, but the duration of the tones were 70 ms and the duration of the speech tokens were 400 ms there were variations in the ISI. However, longer durations enhance ERP components, as would be the result with the longer duration of vowels compared to the tones, and are preferred for these recordings. Therefore, ISI differences would not account for the findings in this experiment.

Natural speech tokens versus synthetic speech tokens. In this experiment, response latencies for P1-N1-P2 components were significantly shorter when recorded with the natural speech tokens than the synthetic speech tokens. There were no significant differences found for P1-N1 or N1-P2 amplitudes. It was initially hypothesized that there would not be any significant differences between responses recorded via the vowel tokens. This hypothesis did not hold true.

A responsible explanation for this finding is that these results are related to the differences in the spectrotemporal characteristics (Agung et al., 2006; Digeser et al., 2009) of the stimuli, especially, differences in the initial onset of the tokens. It has been speculated that natural speech and synthetic speech are processed differently. Supporting evidence for this notion is equivocal. Uppenkamp et al. (2006), via functional magnetic resonance imaging technology, demonstrated that natural tokens and synthetic tokens produced an identical pattern of auditory cortical activation. Benson et al., (2001), however, also used functional magnetic resonance imaging and found differing cortical activations for natural speech versus synthetic speech. Tremblay, Friesen et al. (2003) suggested that cortical processing might be different for natural tokens compared to synthetic tokens, based on response differences from two studies, one using natural speech and the other using synthetic speech.

Male natural and synthetic tokens versus female natural and synthetic tokens. In general, significant differences were not seen across all P1-N1-P2 latency or amplitude components elicited via male voiced tokens (i.e., natural and synthetic male) compared to female voiced tokens (i.e., natural and synthetic voiced) with the only significant finding was statistically shorter P1 latencies for male tokens compared to female tokens. Again, this finding may be related spectral difference between the eliciting tokens. Specifically, the lower frequency content of the male tokens compared to the higher frequency content of the female tokens (see spectrograms, Figures 20 to 23). P1-N1-P2 components are sensitive to stimulus frequency as responses recorded using lower frequency stimuli demonstrate shorter component latencies compared to responses recorded with higher frequency stimuli (Agung et al., 2006; Alain et al., 1997; Antinoro et al., 1969; Jacobson et al., 1992; Sugg & Polich, 1995).

Self-produced tokens versus other speech tokens. In the present study, seven passive listening conditions were employed, two of which, were recorded participant produced /a/ tokens and that same /a/ pitch shifted downward. Responses to these tokens were recorded to examine the effect on component latencies and amplitudes in comparison with the other speech tokens. For the data collected here, no significant differences between self-produced tokens and the other speech tokens were found for P1 latency, P2 latency, P1-N1 amplitude, or N1-P2 amplitude. These findings are consistent with those reported in previous investigations (Heinks-Maldonado et al., 2007; Houde et al., 2002). Significantly shorter N1 latencies were found for responses recorded using self-produced vowels compared to the other natural and synthetic tokens, which is inconsistent with previous studies (Heinks-Maldonado et al., 2005).

As expected, no significant differences were found between the self-produced tokens and the other speech tokens represented in the majority of wave components. This hypothesis was based on the assumption that suppression of cortical activity is the result of motor-to-sensory matching evoked as a consequence of the motor act of vocalization (Frith & Done, 1989; Heinks-Maldonado et al., 2007). Therefore, responses evoked passively with self-produced speech tokens should not differ from the responses collected via "other" passively presented speech tokens.

The finding of significantly shorter N1 latencies for self-produced speech was, however, unexpected. It is likely related to the differences in spectral characteristics of the tokens. One self-produced token was passively presented without an alteration and other token was pitch shifted downward; thus, shifting the spectral content into a lower frequency range. Researchers have reported that N1 amplitudes are larger and latencies shorter when evoked with lower frequency stimuli (Agung et al., 2006; Alain et al., 1997; Antinoro et al., 1969; Jacobson et al.,

1992; Sugg & Polich, 1995). When considering the effect of this frequency shift our finding of shorter N1 latencies for the self-produced speech tokens compared to the other tokens is plausible. Refer to spectrograms presented in Figures 20 to 27.

P1-N1-P2 Component Gender Differences

This investigation was also designed to examine differences between responses from male and female participants. In separating males from females, stimulus or condition effects can be evaluated without the influence of gender (Kudo et al., 2004). Gender differences have been demonstrated across various electrophysiological indices (e.g., early potentials; Beagley et al., 1978; Berghotlz, 1981; Chan et al., 1988; Don et al., 1993); however, findings for late evoked potentials have been equivocal (Dehan & Jerger, 1990; Golgeli et al., 1999; Sabo, Durrant, Curtin, Boston, & Rood, 1992). To follow the trend typically seen it was predicted that wave components recorded from female participants would have significantly shorter latencies and larger amplitudes than male participants. As predicted, P2 latencies were significantly shorter when recorded from females than from males. No other significant gender differences were found for P1 or N1 latencies or component amplitudes.

CHAPTER III: AUDITORY EVENT-RELATED POTENTIALS DURING SPEECH PRODUCTION

Theoretically, incoming auditory inputs are monitored via feedforward and feedback control subsystems (Guenther, 1994, 1995, 2001, 2007). Such monitoring allows for the detection of self-produced errors (Heinks-Maldonado et al., 2005) and the discrimination of internally and externally generated signals (Ford et al., 2002; Frith, 1992). Numerous researchers have documented that self-generated vocalization induces suppression of auditory cortical

responses relative to the same externally generated vocalizations (Ford, Mathalon, Heinks et al., 2001, 2004, Heinks-Maldonado et al., 2005; Houde et al., 2002). It has been suggested that this auditory suppression is the result of motor to sensory integration whereby corollary discharges. That is, the sensory representations of the expected feedback (Sperry, 1950) match with the actual feedback (Blackmore et al., 1998; von Holst, 1954). Under some experimental manipulations, self-generated vocalizations are altered and the speaker does not hear their expected feedback. During these altered conditions, less cortical suppression ensues relative to when the listener hears their own voice unaltered (Houde et al., 2002). The underlying processes of these conditions inducing suppression is not well understood and further work is necessary to understand cortical processing under various altered feedback conditions.

The existence of a corollary discharge is not exclusive to the auditory system. Similar evidence has been found in the visual domain (Welch, 1978) and somatosensory domains (Blackmore, Rees, & Frith, 1998). Additionally, the presence of corollary discharges has been demonstrated across various animal species. It has been suggested that this corollary discharge mechanism provides a method for monitoring self-generated motor functions and allows for the discrimination between self-generated and externally generated stimuli (Sperry, 1950).

Researchers employing neuroimaging techniques have provided support for the occurrence of corollary discharges in the human auditory system by demonstrating the suppression of M100 to self-generated speech. Curio, Neuloh, Numminen, Jousmaki, and Hari (2000), for example, examined such during a speech/replay task. In the speech condition, participants uttered two successive vowels while listening to a random series of two tones. In the replay condition, the same participants listened to the recorded vowel utterances from the speech condition. Interestingly, the M100 responses recorded via the speech task were significantly

delayed in both auditory cortices and reduced in amplitude primarily at the left auditory cortex in comparison to the replay task.

In numerous investigations, Ford and colleagues (Ford, Mathalon, Heinks et al., 2001; Ford, Mathalon, Kalba et al., 2001; Ford et al., 2002; Ford & Mathalon, 2004; Heinks-Maldonado et al., 2005; Heinks-Maldonado et al., 2006; Heinks-Maldonado et al., 2007) have also shown differences in the auditory event related potentials during listening and vocalizations. They demonstrated suppression of the N1 component of the long latency auditory evoked potential to self-generated speech in comparison with N1 responses recording during listening. In light of these findings, they posited that self-vocalization induced N1 suppression (i.e., N1 amplitude reduction) was related to a mismatch between corollary discharge signals and the actual auditory feedback heard when one speaks.

Heinks-Maldonado et al. (2005) noted that N1 responses recorded with altered feedback (i.e., FAF) did not demonstrate the same cortical suppression as those recorded with NAF. They speculated that the introduction of FAF disrupted the forward flow of speech by reducing the amount of cortical matching between the expected feedback and the actual auditory feedback. In other words, auditory cortical suppression is increased when the signal more closely matches the expected feedback. Additionally, based on those findings, they concluded that precision within this forward model is essential for appropriate auditory priming (i.e., N1 suppression) and self-monitoring to take place.

In this investigation the effects of various feedback conditions (i.e., NAF, FAF, DAF 50 ms, and DAF 200 ms) on electrophysiological measures during active vocalizations were also investigated and the following research questions were addressed: Are there differences across conditions of feedback during vocalization? Are there differences between active and passive

conditions? Are there differences between component latency and amplitude between male participants and female participants? Given that the experimental paradigm employed by Heinks-Maldonado et al., (2005) was adapted and utilized for this experiment, hypotheses presented here were formulated primarily based on the findings reported in that study. The experimental hypotheses were as follows: (a) NAF amplitudes will be smaller than FAF and DAF 200 ms amplitudes, (b) DAF 50 ms latencies and amplitudes will be similar to NAF latencies and amplitudes, (c) DAF 50 ms amplitudes will also be smaller than FAF and DAF 200 ms amplitudes, (d) FAF and DAF 200 ms amplitudes and latencies will be similar, (e) active condition amplitudes will be smaller than passive condition amplitudes, and (f) amplitudes will be larger and latencies shorter when recorded from female participants compared to male participants.

Methods

Participants

Apporoval to conduct this research was obtained by the East Carolina University Institutional Review Board prior to data collection or participant recruitment (see Appendix A). Fifteen adult males and 15 adult females whom participated in Experiment 1 served as participants during this investigation. Participants were between the ages of 18 to 30 years, native English speaking, and right-handed. All were recruited from the East Carolina University student body, the College of Allied Health Sciences student body; and/or a research participant pool located in the Department of Communication Sciences and Disorders on a volunteer basis. Informed consent documents were explained, agreed upon, and signed by the participants before testing began (see Appendices B and C).

Apparatus

For all experimental conditions, participants were seated in a comfortable reclining chair placed inside a double walled sound treated audiometric suite (Industrial Acoustics Corporation). This audiometric suite met specifications for permissible ambient noise (American National Standards Institute, 1999). Data acquisition was performed using Compumedics NeuroScan 4.4 software, SynAmps^2 Model 8050 EEG amplifier. A 64-Channel Ag/Ag/Cl Compumedics Quik-Cap was used to record ongoing EEG activity.

The apparatus used for stimulus presentation included a Radio Shack Omnidirectional Dynamic (model 33-3039) microphone. As the participant vocalized the speech signal was initially routed from the microphone to a Mackie Micro Series 1202 mixer. At the level of the mixer the voltage from the speech signal was simultaneously split and sent to a K126 Mini-Voice Onset (VOX) Relay Trigger. The Mini-VOX Relay was connected via USB connection to a Dell Optiplex GX620 computer running Compumedics NeuroScan Stim2 software. A program was written using the Stim2 Gentask module that recognized BYTE inputs from sources such as a mouse click, a keyboard press, or an external button pad press. Once the voltage from the speech signal exceeded 2 volts the VOX relay emulated a mouse click and sent a BYTE signal to Gentask, which in turn, initiated recording. The speech signal was also simultaneously routed to a Yamaha DEQ5 Digital Signal Processor (DSP) where it was altered using transfer functions to emulate the way speech naturally sounds to the speaker. The signal was then routed to a Maico MA53 Audiometer and presented to the participant via 3A Compumedics Neuroscan insert earphones.

Stimulus

N1 and P2 wave components were elicited using self-produced vowel /a/ tokens presented to the listener via NeuroScan ER-3A insert earphones under four feedback conditions:

NAF; a FAF condition where the vocalization was pitch shifted down one half octave; a *long* DAF condition of 200 ms; and a *short* DAF condition of 50 ms. The DSP was set so only the desired alteration was produced. The NAF and FAF conditions were chosen to replicate the previous study conducted by Heinks-Maldonado et al. (2005). The two delay conditions (i.e., 50 and 200 ms) were the same as delays utilized by Kalinowski et al. (1996) and Stuart et al. (2002).

Frequency dependant transfer functions were implemented to emulate natural speech monitoring during test conditions. These transfer functions were derived from several sources. They included Cornelisse, Gagné, and Seewald (1991); Wiener and Ross (1946); and Dean and Martin (2000). Cornelisse, Gagne, and Seewald (1991) calculated the long term average speech spectrum (LTASS) for a microphone placed 30 cm in front of a speaker and at the ear level. These authors demonstrated frequency dependent differences between the LTASS recorded at the reference microphone (i.e., 30 cm in front of the talker) compared to the ear level microphone for males, females, and children. For this investigation the mean for all three groups were used given that both males and females served as participants. Wiener and Ross (1946) calculated sound pressure levels from the ear level to the tympanic membrane using a probe tube microphone apparatus placed next to the tympanic membrane. Transfer functions were calculated by first summing microphones to ear level values from Cornelisse et al. and ear level to tympanic membrane values of Wiener and Ross. Finally, occlusion effect values from Dean and Martin were subtracted. Transfer functions were calculated with the sound source at 0 degrees azimuth. A diagram of how these transfer functions were utilized and a table of the numerical values are presented in Figure 62 and Table 16, respectively.

Stimulus Calibration

To ensure that the spoken /a/ tokens were presented at a consistent intensity (i.e., approximately equal to passive conditions) each participant under went a training session. During this training, each participant was asked to hold an iPhone 3G running the Studio Six Digital SPL Meter application at the level of the Radio Shack microphone. The internal microphone of the iPhone 3G was calibrated within the SPL meter program using a Brüel & Kjær precision sound level meter (type 2231) with a Brüel & Kjær octave band filter (type 1625) attached to a Brüel & Kjær (type 4155) pressure microphone.

Participants were asked to produced the /a/ token at a consistent level around approximately 63 dB (A-weighting) as displayed in the bottom right hand corner of the program. While the participants were vocalizing one of the insert ear phones were coupled to the Brüel & Kjær precision sound level meter (type 2231) with a Brüel & Kjær octave band filter (type 1625) attached to a Brüel & Kjær (type 4144) pressure microphone via a 2 cm^3 (type DB 0138) coupler on the outside of the audiometric suite. The tester monitored the level of the output at the insert earphone to ensure that it was approximately 75 dB SPL. If not, the participant was instructed to increase or decrease vocalization until a level of 75 dB SPL was consistently produced. Once the participant could consistently produce the /a/ token approximately 10 to 15 times the training session was halted and testing began.

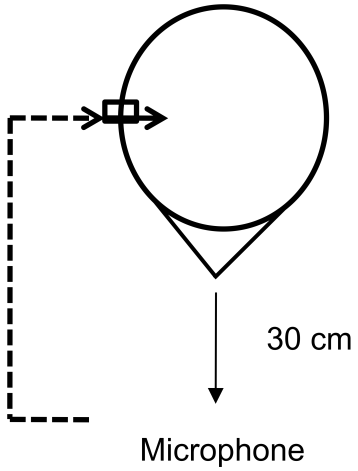


Figure 62. Schematic of implemented transfer functions. The circle represents the participants' head facing a microphone placed at a distance of 30 cm. The dashed arrow represents transfer functions used from Cornelisse et al., (1991). The solid black arrow represents the transfer functions from Wiener & Ross (1946). The outlined box represents an insert earphone.

Table 16.

Numerical Values of Implemented Transfer Functions.

1/3 Octave Band Levels (Hz)	30 cm to ear level Mean (dB SPL)[a]	Ear level to Eardrum (dB SPL)[b]	Deep Insertion of the Insert headphone (dB)[c]	Implemented Transfer Function
200	4.8	0		4.8
250	5.9	0.2	9	-2.9
315	5.9	0.2		6.1
400	4.9	0.2		5.1
500	5.0	0.2	6	-3.8
630	5.1	0.5		5.6
800	5.6	0.5		6.1
1000	5.5	1.0	1	5.5
1250	-0.5	2.0		1.5
1600	-1.0	3.0		2.0
2000	1.0	5.0		6.0
2500	1.0	6.5		7.5
3150	-3.0	10.5		7.5
4000	-6.9	10.0		3.1
5000	-6.0	6.0		0
6300	-5.0	2.5		-2.5
8000	-9.5	4.0		-5.5

Note. All values were rounded to the nearest whole number. [a]Values from Cornelisse et al. (1991); [b]Wiener & Ross (1946); [c]Dean & Martin (2000). The implemented transfer function was calculated as (a+b)-c.

Procedure

Prior to testing, the circumference of the participant's head was measured to ensure appropriate cap sizing. The Compumedics NeuroScan Quik-Caps come in various sizes and styles. Three sizes were available for use during this experiment. Head circumference measurements of 50 to 55 cm, 56 to 59 cm, and 60 to 65 cm fit cap sizes small, medium, and large, respectively. Electrodes are placed on the Quik-Cap according to the 10/20 International Electrode System (Jasper, 1958).

During this experiment all participants were asked to sit and vocalize the vowel /a/ every 1 to 2 seconds into the microphone while their speech token was presented via insert earphones under the various conditions. This microphone was positioned at 0 degrees azimuth from the participant's mouth at a distance of 30 cm with a Radio Shack microphone boom stand. This microphone positioning was at an appropriate distance from the participant so vocalizations were detected and testing not disrupted. They were also asked to remain alert and minimize movements especially head and neck movements while the stimulus was presented. Frequent breaks were given throughout testing to ensure participant comfort and avoid listener fatigue. Stimulus conditions were counter balanced according to a diagram balanced Latin squares design (Wagenaar, 1969) and bilaterally presented using insert earphones. Insert earphones were placed in the participant's ears using methodology presented by Dean and Martin (2000).

EEG activity was recorded from eleven electrode sites placed about the participant's head (i.e., F3, Fz, F4, C3, Cz, C4, T4, T3, Pz, M1, and M2). While recording, the nasion served as the reference and Fpz as the common ground. Vertical eye movements were monitored using electrodes placed vertically above and below the left eye. Electrode impedances were maintained at or below 5000 Ω. EEG activity was amplified 3000 times and analog filtered (i.e., 1-30 Hz).

Analog-to-digital conversion at a rate of 500 /s for all channels was performed with a PC based NeuroScan system and SynAmps^2 24 bit amplifier. All EEG activity was saved as continuous files for offline averaging and digital filtering.

Off-line Waveform Analyses

The same waveform analysis was employed as in Experiment 1. All continuous files were epoched to stimulus onset in a time locked window from -100 to +500 ms. This response window was selected to provide an adequate time window for visualization of all pertinent waveform components (Martin & Stapells, 2005; Sharma, Marsh, & Dorman, 2000). A pre-stimulus interval of -100 was used as the baseline correction for all recording channels. Epochs were then digitally filtered (i.e., 1-40 Hz, 24 dB/octave) and re-referenced to linked mastoids (M1 and M2). Ocular movement artifacts were digitally removed from the epochs (Semlitsch, Anderer, Schuster, & Presslich, 1986). Epochs containing artifacts exceeding +/- 50 μV were rejected from averaging. P1-N1-P2 waveform components were replicated between two trials. Replication was defined as matching waveform components within 25 ms (O'Brien & Stuart, 2001). Replications were then averaged creating one waveform for each condition for a total of four waveforms for each individual participant.

Electrophysiological Waveform Analyses

When comparing waveforms collected during passive listening conditions and active speaking conditions, one of the most notable differences, was the considerable difference in the overall waveform morphology. P1-N1-P2 waveform components were reliably and consistently present in responses recorded during passive listening conditions. Responses recorded via active speaking conditions presented with a degraded P1 waveform component in almost all participants. Fortunately, N1 and P2 waveform components remained intact and could be reliably

indentified. Reporting only the N1 and P2 latency and amplitude values from the waveforms collected under active speaking conditions is consistent with other investigators (Ford & Mathalon, 2004; Heinks-Maldonado et al., 2005). Waveform components, N1 and P2 were evaluated in terms of latency and amplitude measures. N1 was defined as the largest negative deflection and P2 was defined as the following positive peak, between N1 and 300 ms (Ford, Mathalon, Heinks et al., 2001).

Based on values used by Kraus et al. (1993) N1 and P2 were considered present if amplitudes are equal to or exceed 0.5 μV. Amplitudes were measured from component peak to the adjacent trough. Waveforms components were also considered present if greater responses were visualized from fronto-central electrode sites in comparison to the parietal electrode. Further, polarity inversion at the mastoid sites was also used to determine response presence (Vaughan & Ritter, 1970). However, this polarity inversion was not required for the response to be considered present (Martin & Stapells, 2005; Tremblay, Friesen et al., 2003).

Results

Grand average waveforms for NAF, FAF, DAF 50 ms, and DAF 200 ms conditions recorded across all electrode sites are presented respectively for male participants in Figures (63, 67, 71, and 75) and for female participants in Figures (64, 68, 72, and 76). Individual participant waveforms for NAF, FAF, DAF 50 ms, and DAF 200 ms experimental conditions are presented in Figures 65, 69, 73, and 77, respectively. Collected waveforms for male and female participants are presented in left and right panels, respectively. The grand average waveform is shown as the thickest black line. The grand average waveforms are also presented individually in Figures 66, 70, 74, and 78. Waveforms presented below were recorded from the Cz electrode site.

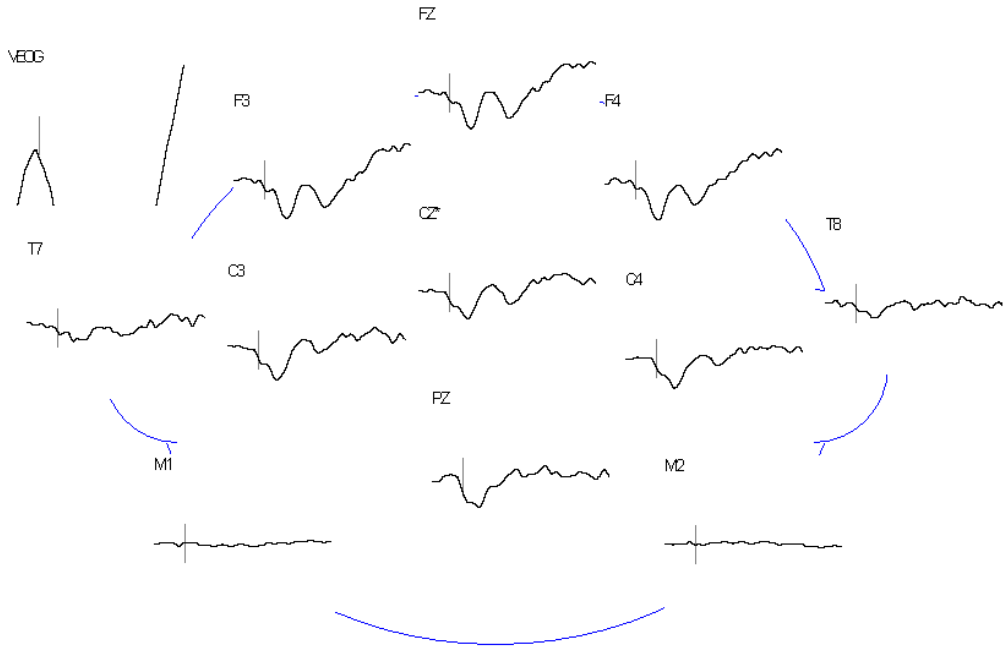


Figure 63. Grand average waveforms recorded across all electrode sites for male participants elicited via active self-produced /a/ tokens presented under NAF.

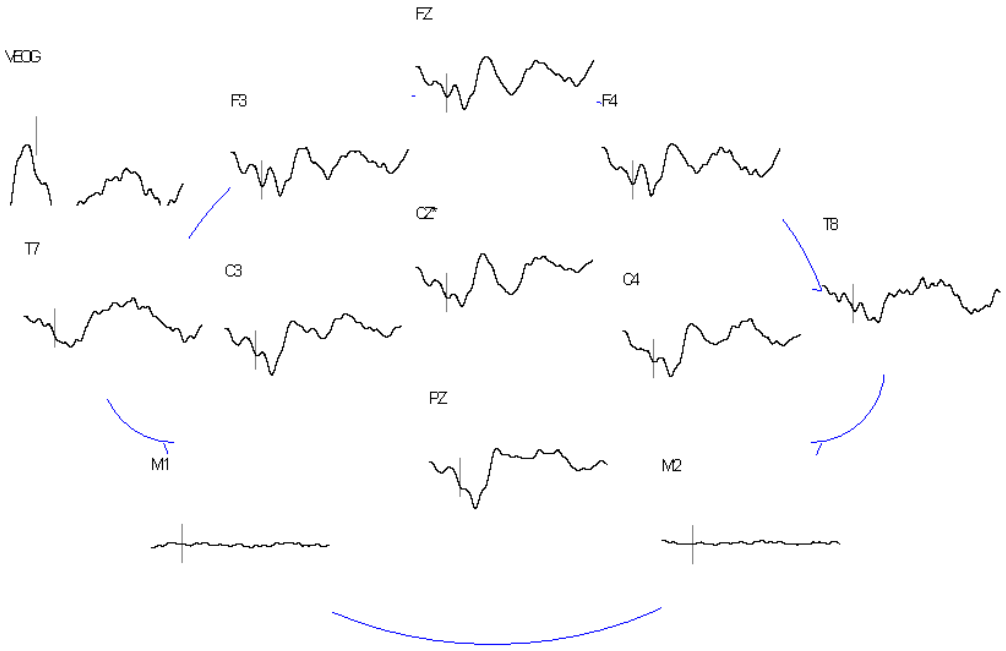


Figure 64. Grand average waveforms recorded across all electrode sites for female participants elicited via active self-produced /a/ tokens presented under NAF.

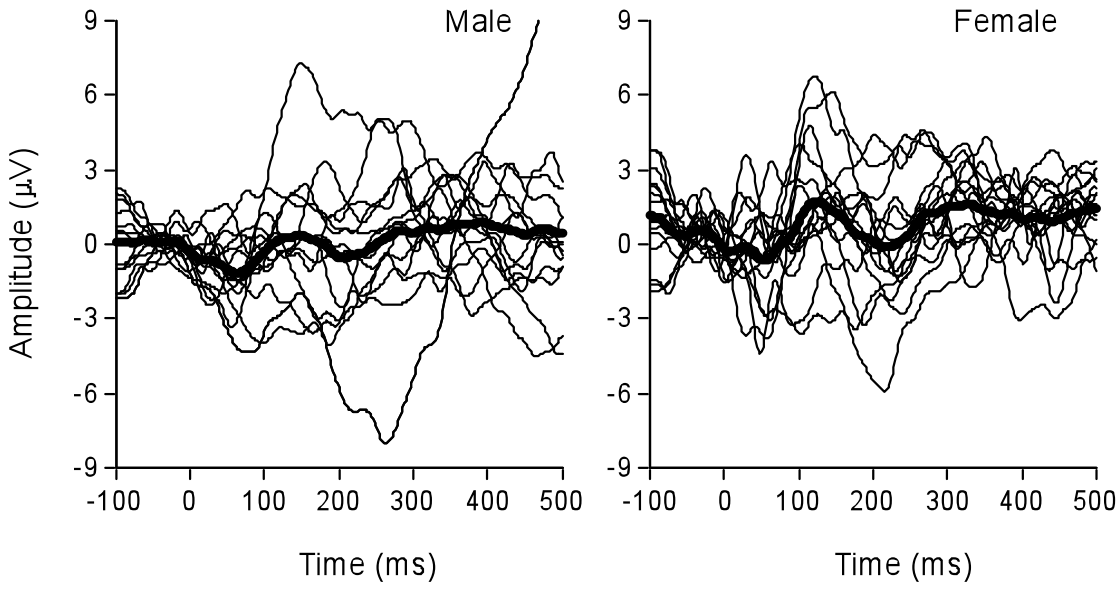


Figure 65. Individual participant waveforms recorded at Cz for male and female participants elicited via active self-produced /a/ presented under NAF showing amplitude (μV) as a function of time (ms). The grand average waveforms are displayed as the thickest black line.

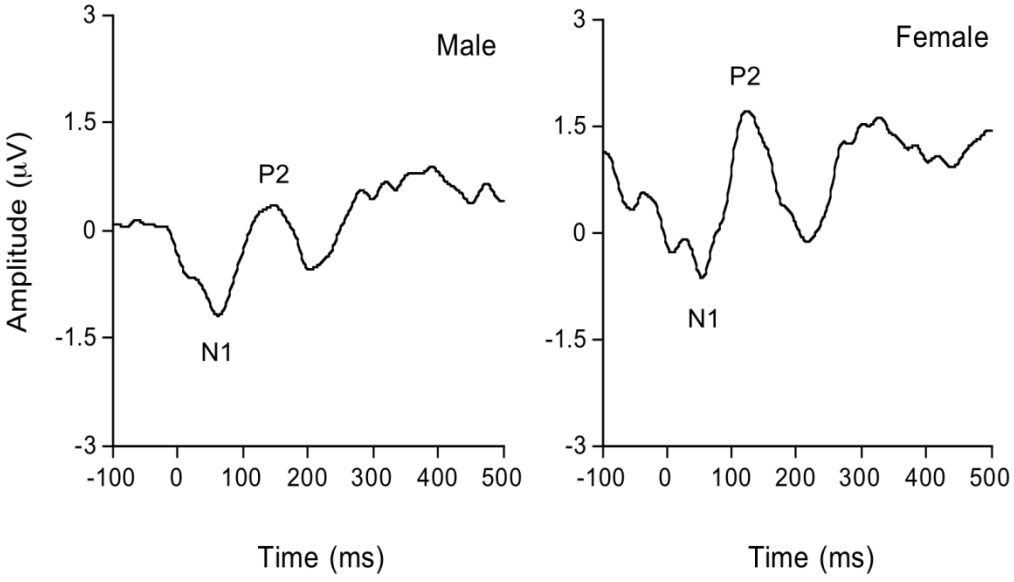


Figure 66. Grand average waveforms recorded at Cz for male and female participants elicited via self-produced /a/ tokens presented under NAF showing amplitude (μV) as a function of time (ms).

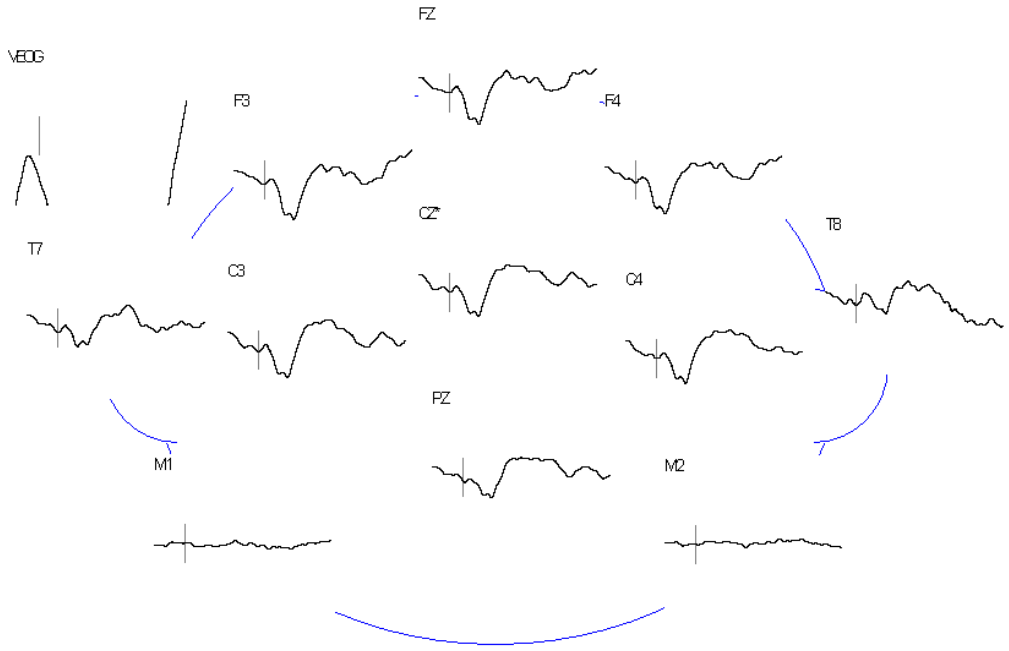


Figure 67. Grand average waveforms recorded across all electrode sites for male participants elicited via active self-produced /a/ tokens presented under FAF.

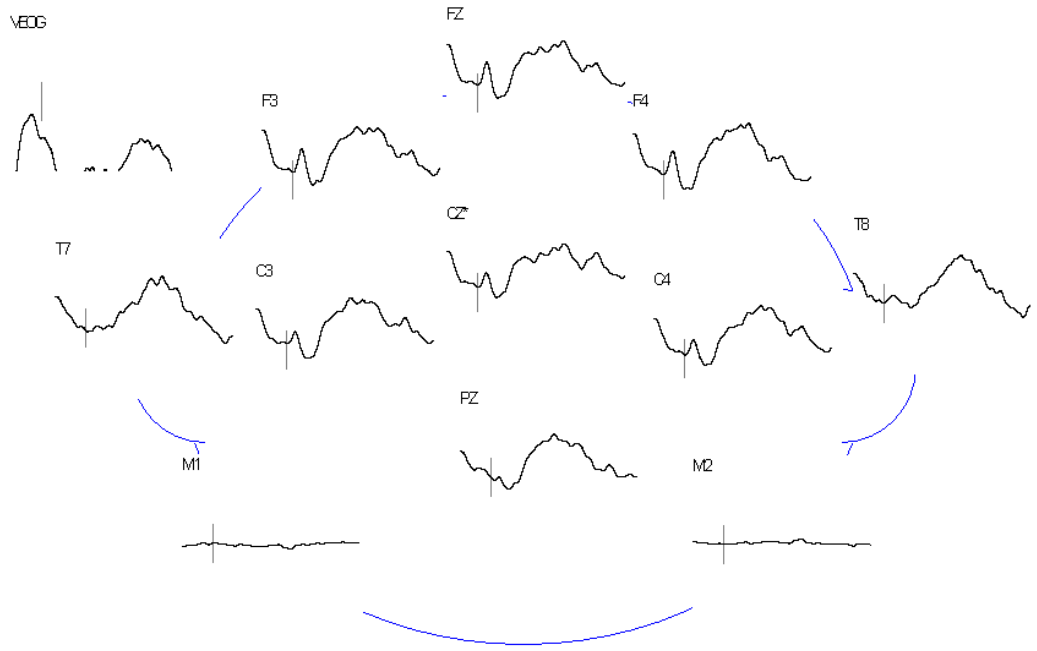


Figure 68. Grand average waveforms recorded across all electrode sites for female participants elicited via active self-produced /a/ tokens presented under FAF.

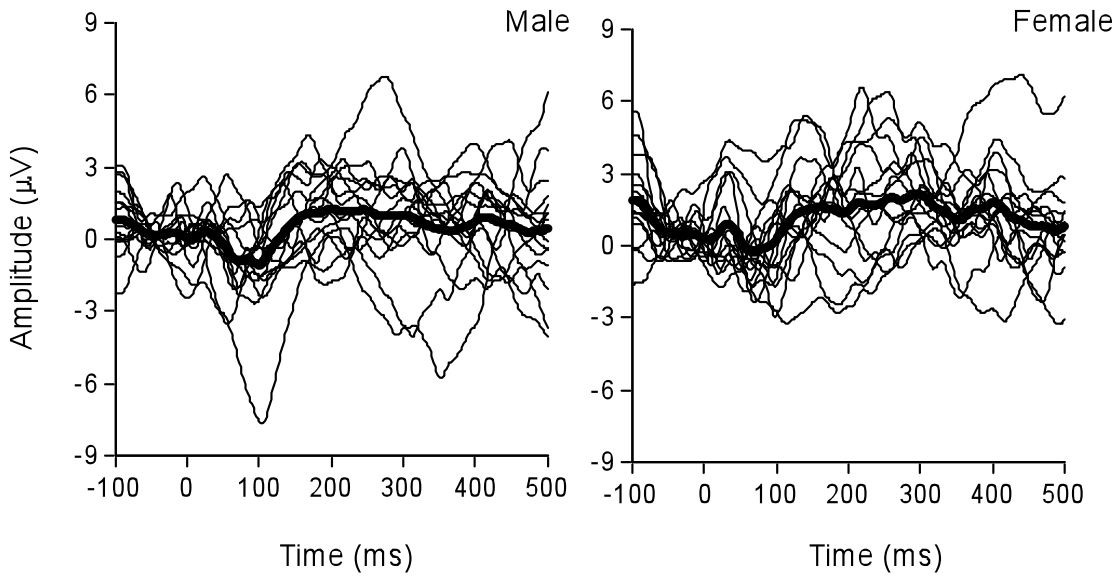


Figure 69. Individual participant waveforms recorded at Cz for male and female participants elicited via active self-produced /a/ tokens presented under FAF showing amplitude (μV) as a function of time (ms). The grand average waveforms are displayed as the thickest black line.

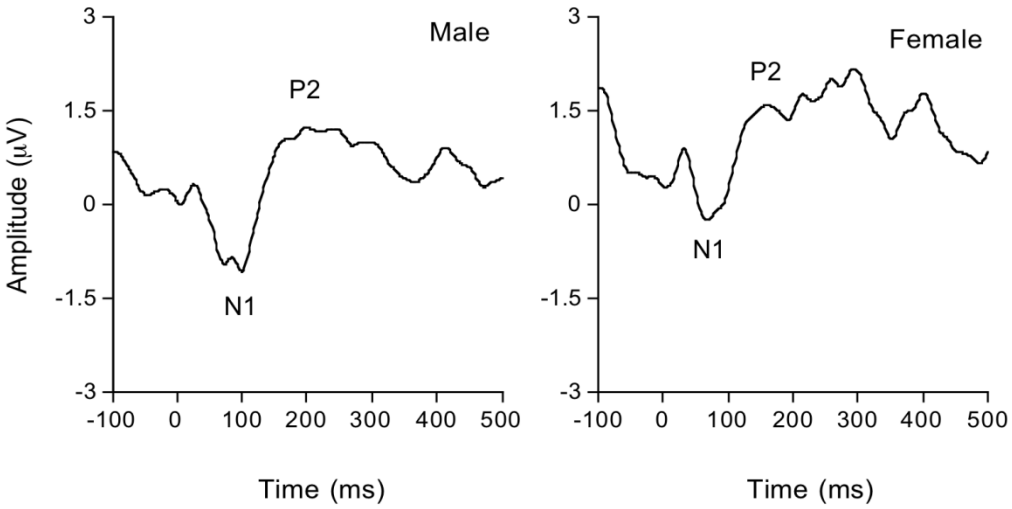


Figure 70. Grand average waveforms recorded at Cz for male and female participants elicited via self-produced /a/ tokens presented under FAF showing amplitude (μV) as a function of time (ms).

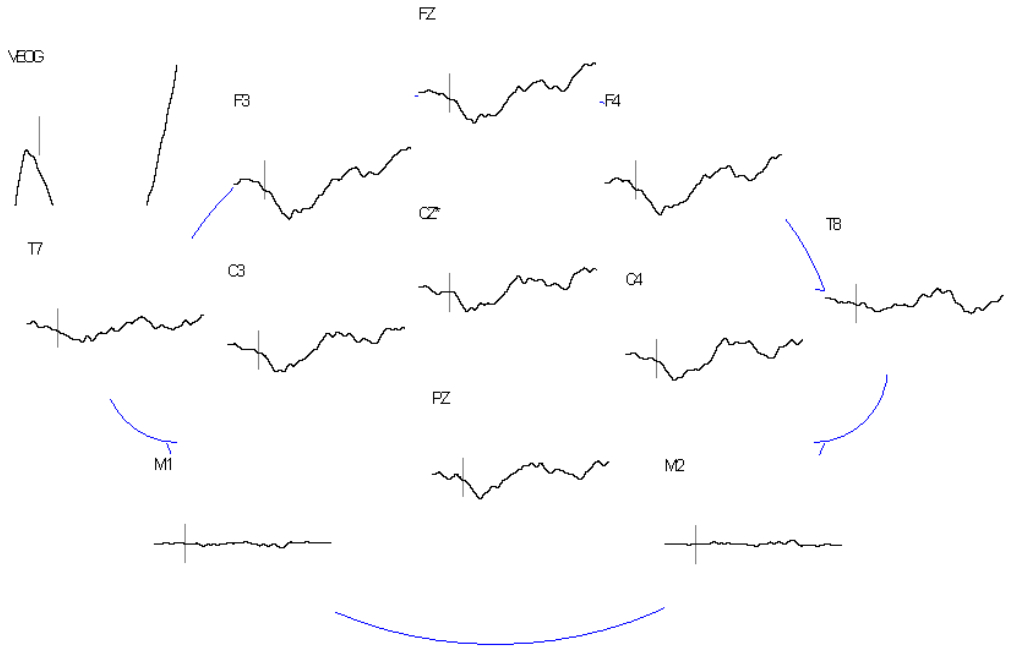


Figure 71. Grand average waveforms recorded across all electrode sites for male participants elicited via active self-produced /a/ tokens presented under DAF 50 ms.

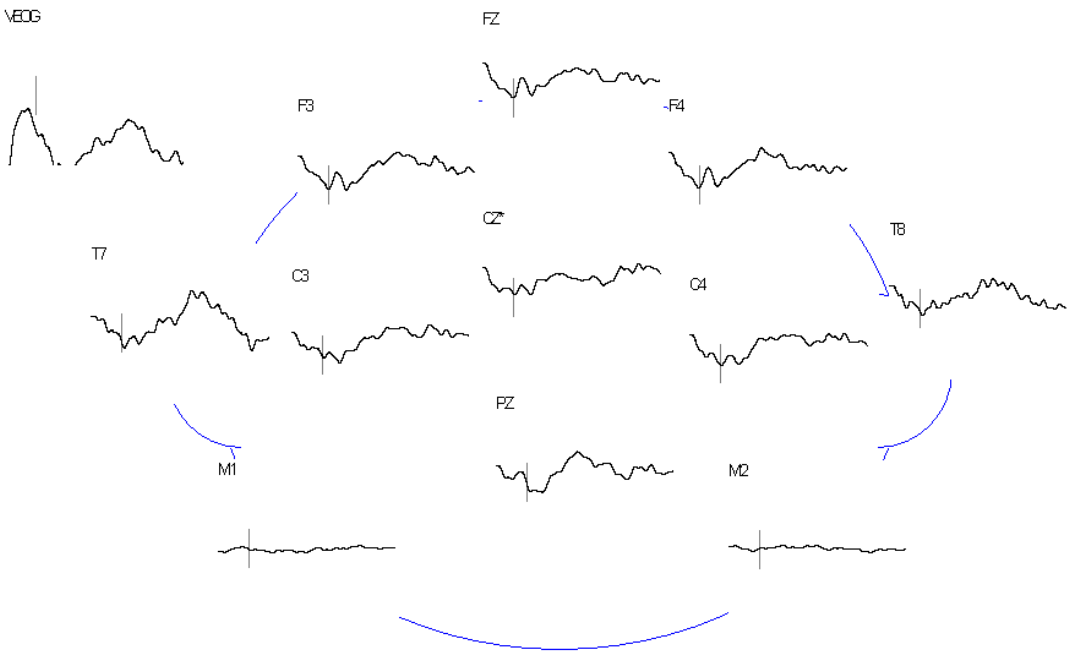


Figure 72. Grand average waveforms recorded across all electrode sites for female participants elicited via active self-produced /a/ tokens presented under DAF 50 ms.

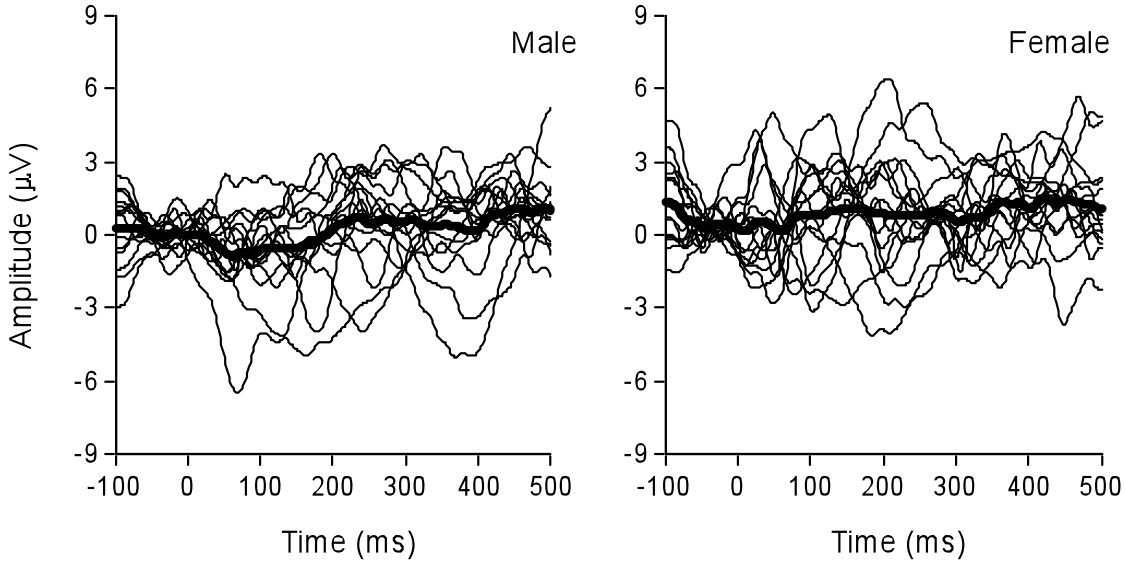


Figure 73. Individual participant waveforms recorded at Cz for male and female participants elicited via active self-produced /a/ tokens presented under DAF 50 ms showing amplitude (μV) as a function of time (ms). The grand average waveforms are displayed as the thickest black line.

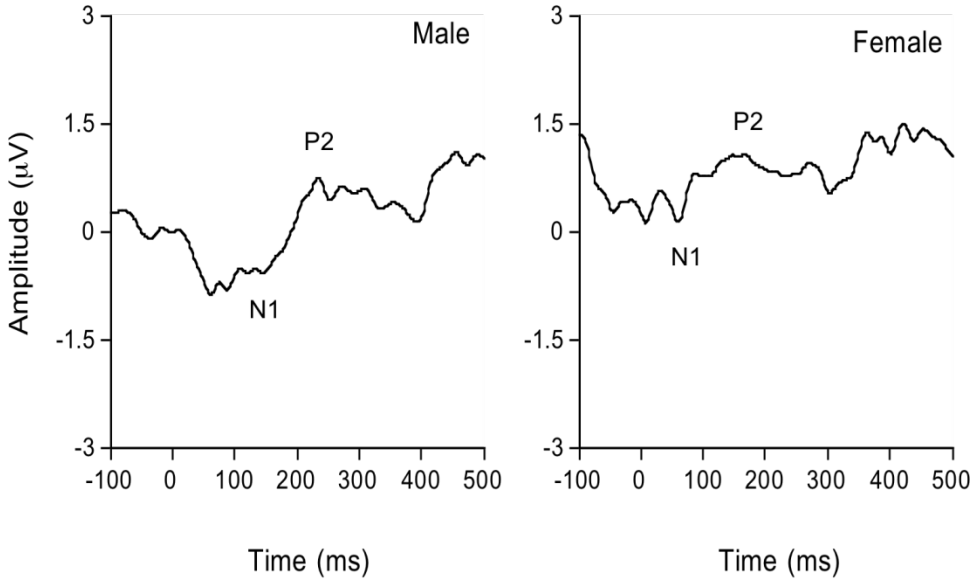


Figure 74. Grand average waveforms recorded at Cz for male and female participants elicited via DAF 50 ms showing amplitude (μV) as a function of time (ms).

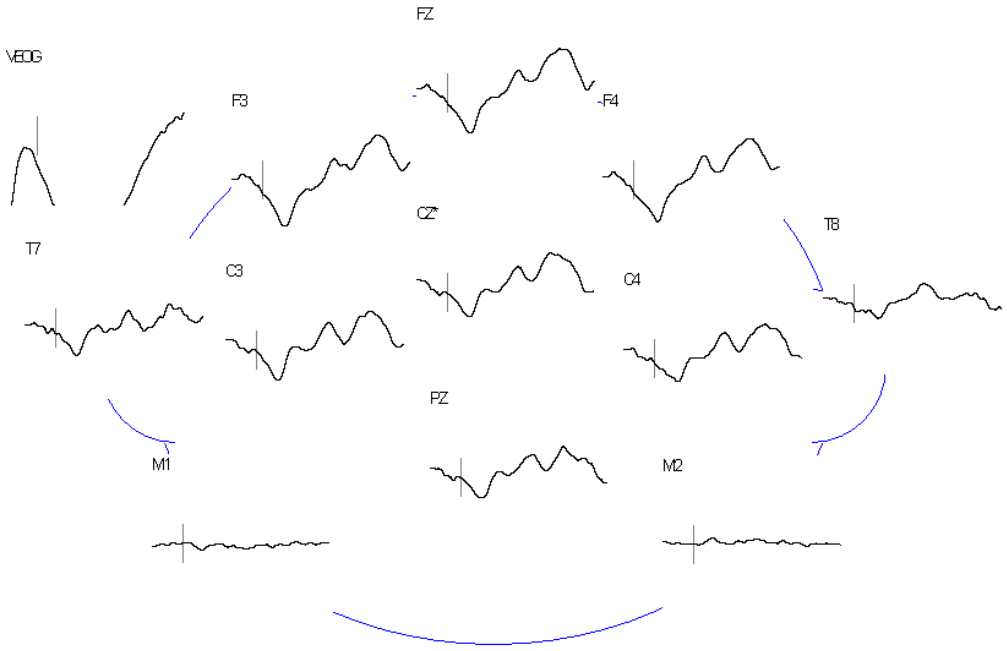


Figure 75. Grand average waveforms recorded across all electrode sites for male participants elicited via active self-produced /a/ tokens presented under DAF 200 ms.

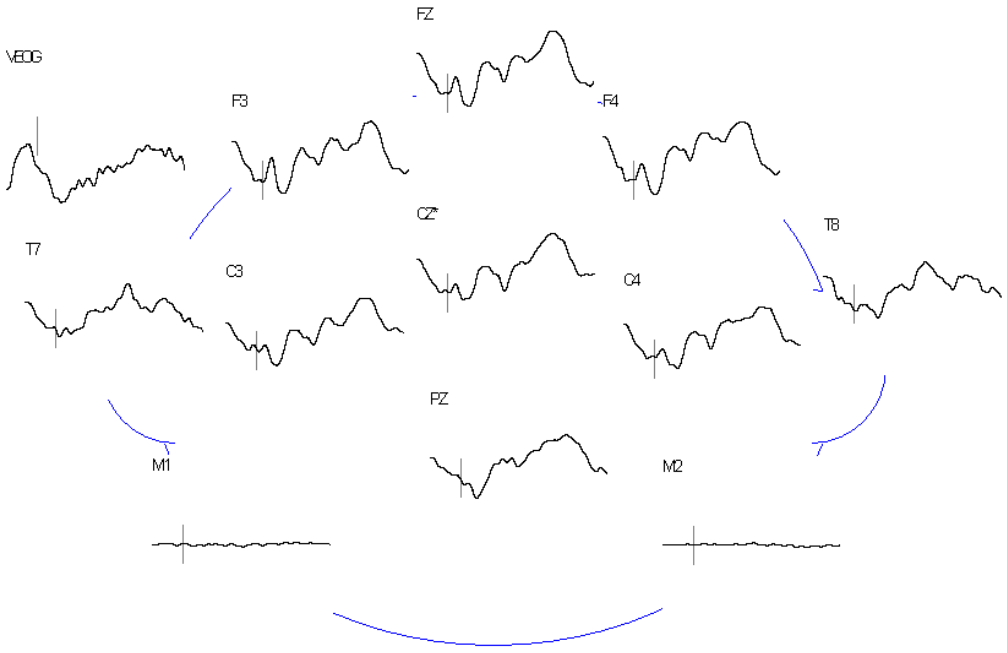


Figure 76. Grand average waveforms recorded across all electrode sites for female participants elicited via active self-produced /a/ tokens presented under DAF 200 ms.

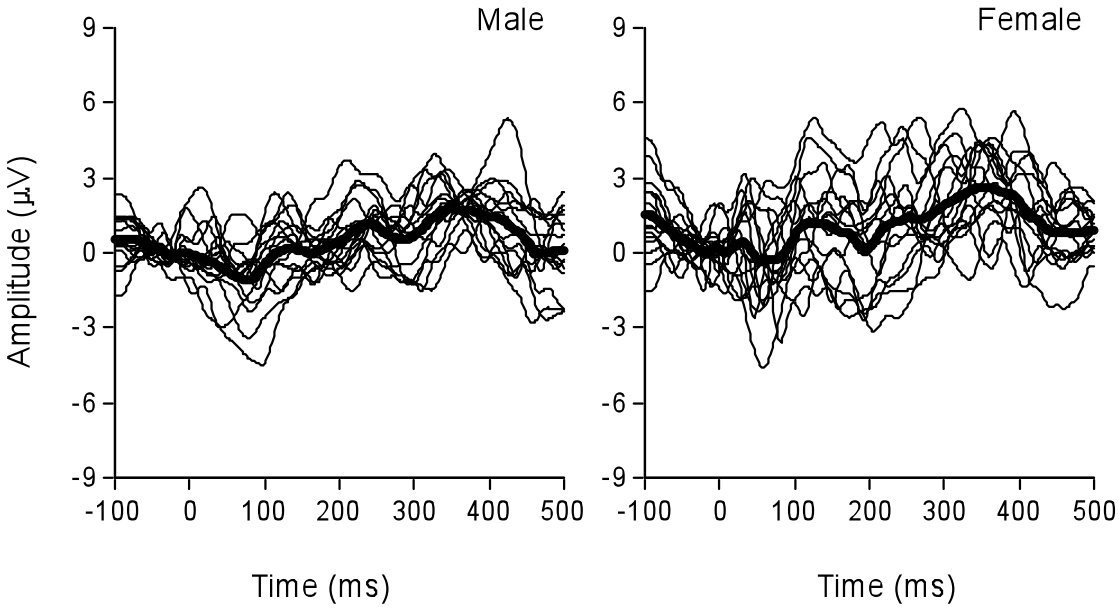


Figure 77. Individual participant waveforms recorded at Cz for male and female participants elicited via active self-produced /a/ presented under DAF 200 ms showing amplitude (μV) as a function of time (ms). The grand average waveforms are displayed as the thickest black line.

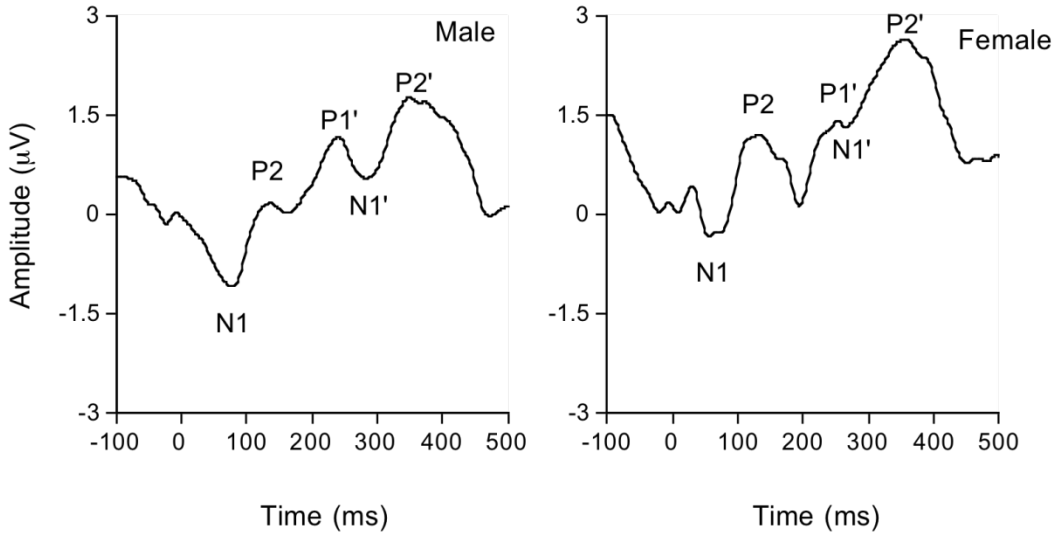


Figure 78. Grand average waveforms recorded at Cz for male and female participants elicited via DAF 200 ms showing amplitude (μV) as a function of time (ms).

Auditory Feedback Conditions

N1 latency. Mean latencies and standard deviations for the N1 wave component, as a function of condition and gender are numerically presented in Table 17. A two-factor mixed ANOVA was utilized to examine the effect of stimulus condition and gender on N1 latency. The ANOVA summary is presented in Table 18. Significant main effects of condition and gender and a significant interaction of condition by gender were found. Figure 79 illustrates the interaction of gender and stimulus conditions.

To further evaluate the source of the interaction post hoc independent and paired *t*-tests were utilized. Independent *t*-tests examined differences across gender. For the DAF 50 ms condition, N1 latencies were significantly shorter in female participants $t(28) = 4.78$, $p < .0001$ than male participants. No other significant N1 latency differences were found between the male and female participants.

Paired *t*-tests were conducted to separately evaluate N1 latency differences between the altered feedback conditions for each gender. For male participants, significant N1 latency differences were found between the following conditions: NAF and DAF 50 ms, $t(14) = 8.32$, $p< .0001$; FAF and DAF 50 ms, $t(14) = 6.25$, $p < .0001$; FAF and DAF 200 ms, $t(14) = 3.22$, $p = .006$; and DAF 50 ms and DAF 200 ms $t(14) = 9.50$, $p < .0001$. These results indicated that N1 latencies were significantly shorter for: NAF, FAF, and DAF 200 ms compared to DAF 50 ms, as well as DAF 200 ms compared to FAF. For female participants, significant N1 latency differences were found between the NAF and DAF 50 ms, $t(14) = 2.59$, $p = .022$. This result is consistent with significantly shorter N1 latencies for the NAF condition compared to the DAF 50 ms condition.

Table 17.

N1 Mean Peak Latencies and Standard Deviations as a Function of Condition and Gender.

	Latency (ms)	
Stimuli	Male	Female
NAF	82.7	66.3
	(23.0)	(23.2)
FAF	95.2	85.9
	(20.8)	(25.4)
DAF 50 ms	159.2	96.5
	(35.7)	(36.2)
DAF 200 ms	76.1	79.9
	(15.7)	(28.4)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

Table 18.

Summary of Two-Factor ANOVA Comparing Differences Between N1 Mean Peak Latency (ms) as a Function of Condition (i.e., NAF, FAF, DAF 50 ms, and DAF 200 ms) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	53697.97	3	17899.32	30.88	<.0001*	.52
Gender	13440.83	1	13440.83	11.71	.002*	.23
Gender X Condition	18787.57	3	6262.52	10.81	<.0001*	.28

Note. * significant at $p < .05$.

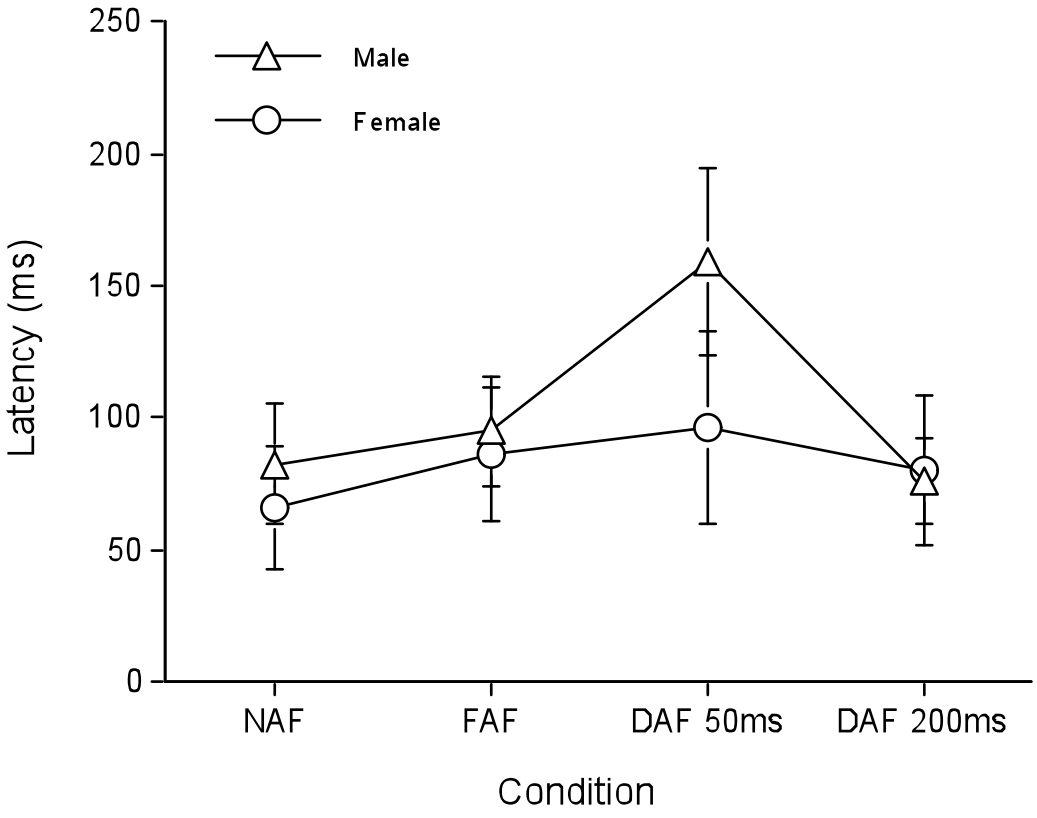


Figure 79. N1 mean peak latencies as a function of condition and gender. Error bars represent +/- one *SD* of the mean.

P2 latency. Mean P2 latencies and standard deviations are numerically presented in Table 19 as a function of condition and gender. A two-factor mixed ANOVA was utilized to examine the effect of stimulus condition and gender on P2 latency (see Table 20). Statistically significant main effects found for condition and gender as well as a significant condition by gender interaction (see Figure 80). To further evaluate the source of the interaction post-hoc independent and paired *t*-tests were utilized. Independent *t*-tests were used to examine P2 latency differences across gender. This result indicated that P2 latencies recorded from the female participants were significantly shorter than those recorded from male participants for the FAF condition, $t(28) = 2.41$, $p = .023$ and the DAF 50 ms condition, $t(28) = 3.66$, $p < .0001$. P2 latencies were not significantly different between male and female participants for the NAF condition $t(28) = 0.23$, $p = 0.80$ and the DAF 200 ms $t(28) = 0.60$, $p = 0.55$ condition.

Paired *t*-tests were separately conducted for male and female groups to evaluate P2 latencies differences between feedback conditions for each gender. For the male participants, significant P2 latency differences were found between the following conditions: NAF and FAF $t(14) = 3.20$, $p = .006$; NAF and DAF 50 ms $t(14) = 6.24$, $p < .0001$; FAF and DAF 50 ms, $t(14) = 3.50$, $p = .004$; FAF and DAF 200 ms, $t(14) = 4.36$, $p = .001$; and DAF 50 ms and DAF 200 ms, $t(14) = 6.19$, $p < .0001$. These results indicate significantly shorter P2 latencies for the following pairs: NAF compared to FAF, NAF compared to DAF 50 ms, FAF compared to DAF 50 ms, DAF 200 ms compared to FAF, and DAF 200 ms compared to DAF 50 ms. For the female group, a significant P2 latency difference was found between the DAF 50 ms and DAF 200 ms conditions, $t(14) = .369$, $p = .013$. This result is consistent with significantly shorter P2 latencies for the DAF 200 ms condition compared to the DAF 50 ms condition.

Table 19.

P2 Mean Peak Latencies and Standard Deviations as a Function of Condition and Gender.

	Latency (ms)	
Stimuli	Male	Female
NAF	140.5	137.3
	(24.7)	(41.5)
FAF	173.2	148.0
	(29.6)	(27.5)
DAF 50 ms	221.2	164.8
	(37.8)	(46.2)
DAF 200 ms	136.9	131.7
	(29.2)	(16.7)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

Table 20.

Summary of Two-Factor ANOVA Comparing Differences Between P2 Mean Peak Latency (ms) as a Function of Condition (i.e., NAF, FAF, DAF 50 ms, and DAF 200 ms) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	64464.63	3	21488.21	23.26	<.0001*	.45
Gender	15187.50	3	15187.50	9.73	.004*	.15
Gender X Condition	13712.10	1	4570.70	4.95	.003*	.26

Note. * significant at $p < .05$.

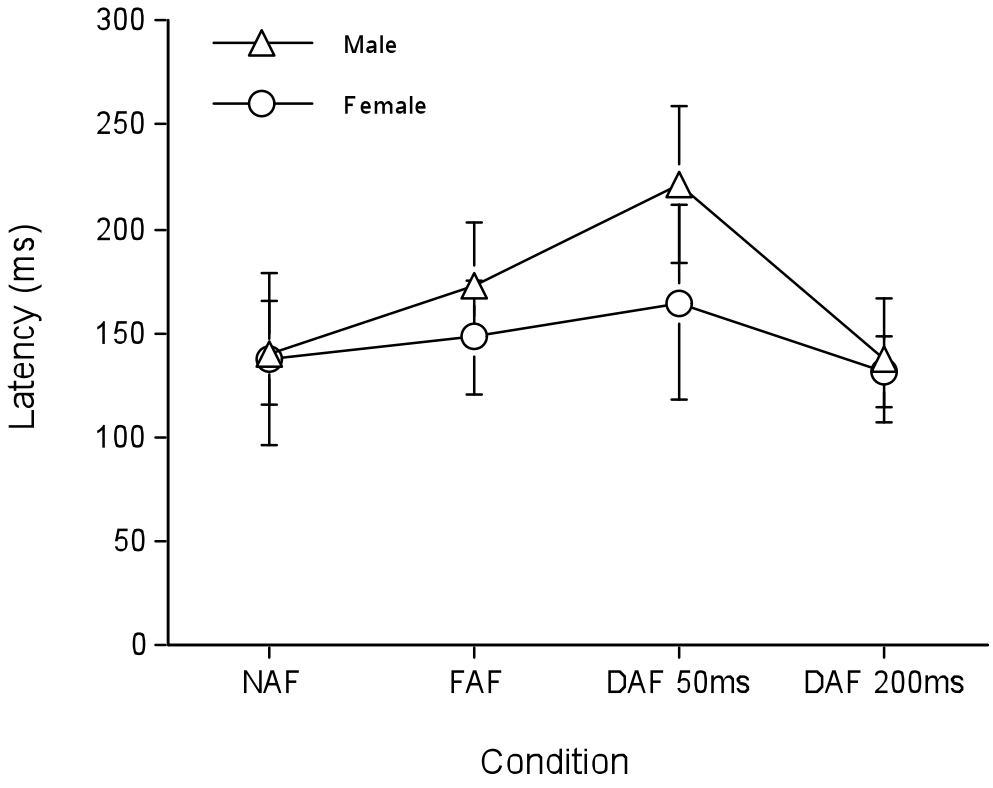


Figure 80. P2 mean peak latencies as a function of condition and gender. Error bars represent +/- one *SD* of the mean.

N1-P2 complex amplitude. Mean N1-P2 peak amplitudes and standard deviations for the N1-P2 complex are presented as a function of stimulus condition and gender in Table 21. A two-factor mixed ANOVA was performed to examine the effect of stimulus condition and gender on N1-P2 waveform amplitude. The ANOVA summaries for N1-P2 wave component analyses are presented in Table 22. As evident in this table, no significant main effects or interactions were found. The grand mean for N1-P2 amplitude collapsed across conditions is 3.02 μV (95% CI 2.45-3.60).

Passive Listening Versus Active Speaking

N1 latency. One focus of this investigation was to look at the differences between auditory evoked potentials elicited during passive listening and actively producing a vowel token. To do so, evoked responses recorded with the self-produced /a/ tokens and the frequency shifted self-produced /a/ tokens during Experiment 1 were compared to the corresponding NAF and FAF conditions in the active speaking condition. N1 latency means and standard deviations are presented in Table 23 as a function of condition and gender and graphically presented in Figure 81. A three-factor mixed ANOVA was utilized to examine the effect of stimulus condition (i.e., active and passive), feedback (i.e., NAF and FAF), and gender. ANOVA results are summarized in Table 24. Significant main effects of condition, feedback, and gender were found. A significant interaction of condition by feedback was also found (see Figure 82). The grand average waveforms for the passive NAF and the active NAF and the passive FAF and the active FAF conditions are presented in Figures 83 and 84, respectively.

Three single-*df* contrasts were utilized to examine the source of this interaction. Statistically significant differences were found between the NAF and FAF in the active condition ($p = .0014$; cf. 74.5 ms vs. 90.6 ms). The passive (i.e., NAF and FAF) conditions and the active

Table 21.

N1-P2 Mean Amplitudes and Standard Deviations as a Function of Condition and Gender.

	Amplitude (μV)	
Stimuli	Male	Female
NAF	2.6	3.9
	(2.3)	(2.4)
FAF	3.3	3.4
	(2.0)	(1.6)
DAF 50 ms	2.7	3.3
	(1.6)	(1.7)
DAF 200 ms	2.1	3.0
	(1.1)	(2.2)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

Table 22.

Summary of Two-Factor ANOVA Comparing Differences Between N1-P2 Mean Peak Amplitude (μV) as a Function of Condition (i.e., NAF, FAF, DAF 50 ms, and DAF 200 ms) and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	12.92	3	4.31	2.43	.071	.08
Gender	16.60	1	16.60	1.76	.20	.06
Gender X Condition	5.73	3	1.91	1.08	.36	.04

Note. * significant at $p < .05$.

Table 23.

N1 Mean Peak Latencies and Standard Deviations as a Function of Condition (i.e., Passive vs. Active), Feedback (i.e., NAF vs. FAF), and Gender (i.e., Male vs. Female).

		Latency (ms)	
Condition	Feedback	Male	Female
Passive	NAF	123.7	109.7
		(16.3)	(13.9)
	FAF	122.1	113.1
		(15.5)	(17.2)
Active	NAF	82.7	66.8
		(23.0)	(23.2)
	FAF	95.2	85.9
		(20.8)	(25.4)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

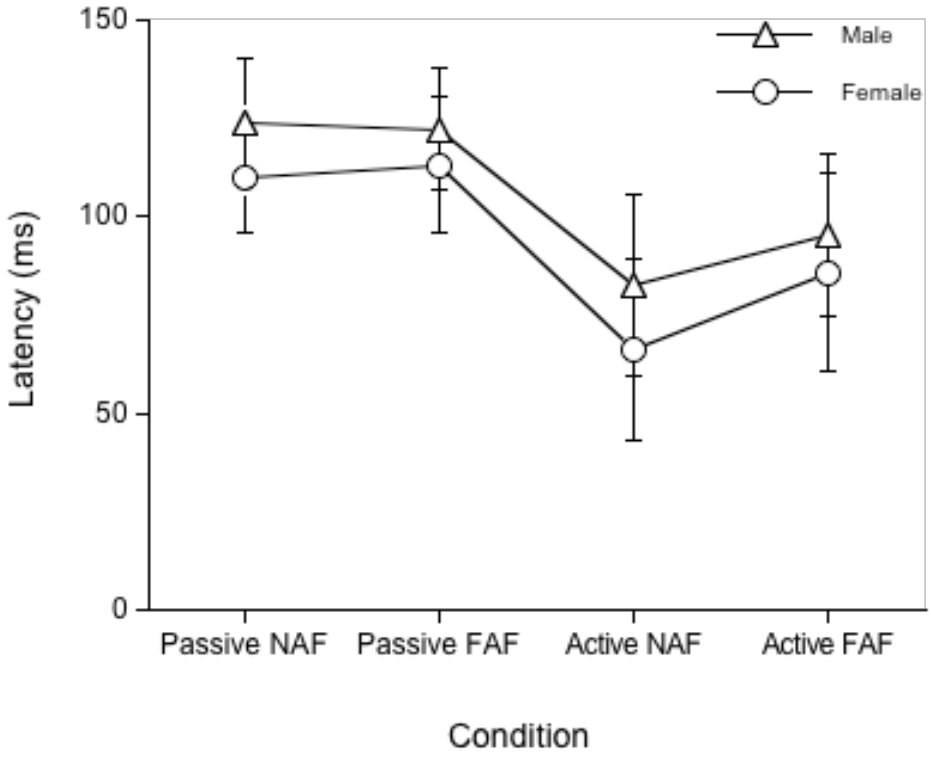


Figure 81. N1 mean peak latencies as a function of condition, feedback, and gender. Error bars represent +/- one *SD* of the mean.

Table 24.

Summary of Three-Factor mixed ANOVA Comparing Differences Between N1 Mean Peak Latency (ms) as a Function of Condition (i.e., Active vs. Passive), Feedback (i.e., NAF vs. FAF), and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	36053.33	1	36053.33	83.43	<.0001*	.75
Feedback	2150.53	1	2150.53	7.39	.011*	.21
Gender	4465.20	1	4465.20	8.38	.007*	.23
Condition X Gender	13.33	1	13.33	.03	.86	.00
Feedback X Gender	270.00	1	270.00	.93	.34	.03
Condition X Feedback	1732.80	1	1732.80	5.60	.025*	.17
Condition X Feedback X Gender	8.53	1	8.53	.03	.87	.00

Note. * significant at $p < .05$.

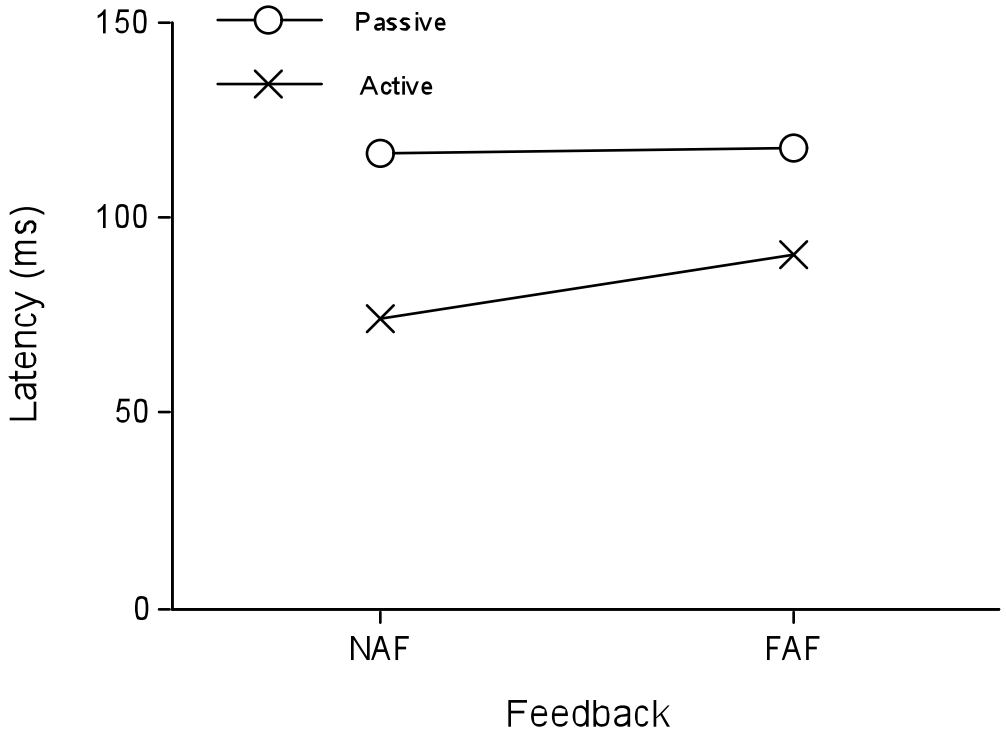


Figure 82. Interaction plot of N1 mean peak latency collapsed across gender as a function of condition (i.e., active vs. passive) and feedback (i.e., NAF vs. FAF).

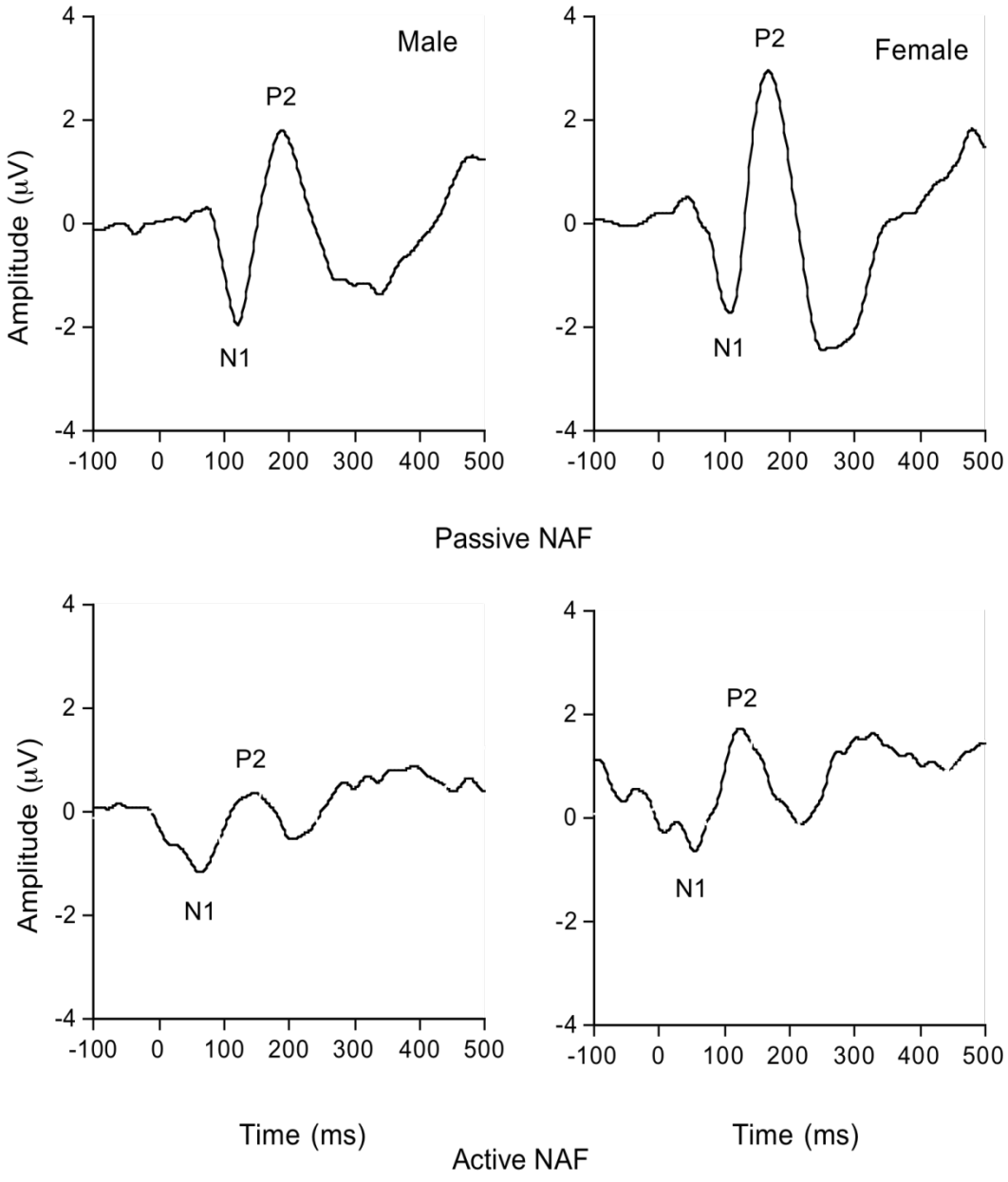


Figure 83. Grand average waveform for the passive NAF condition compared to the grand average waveform for the active NAF condition.

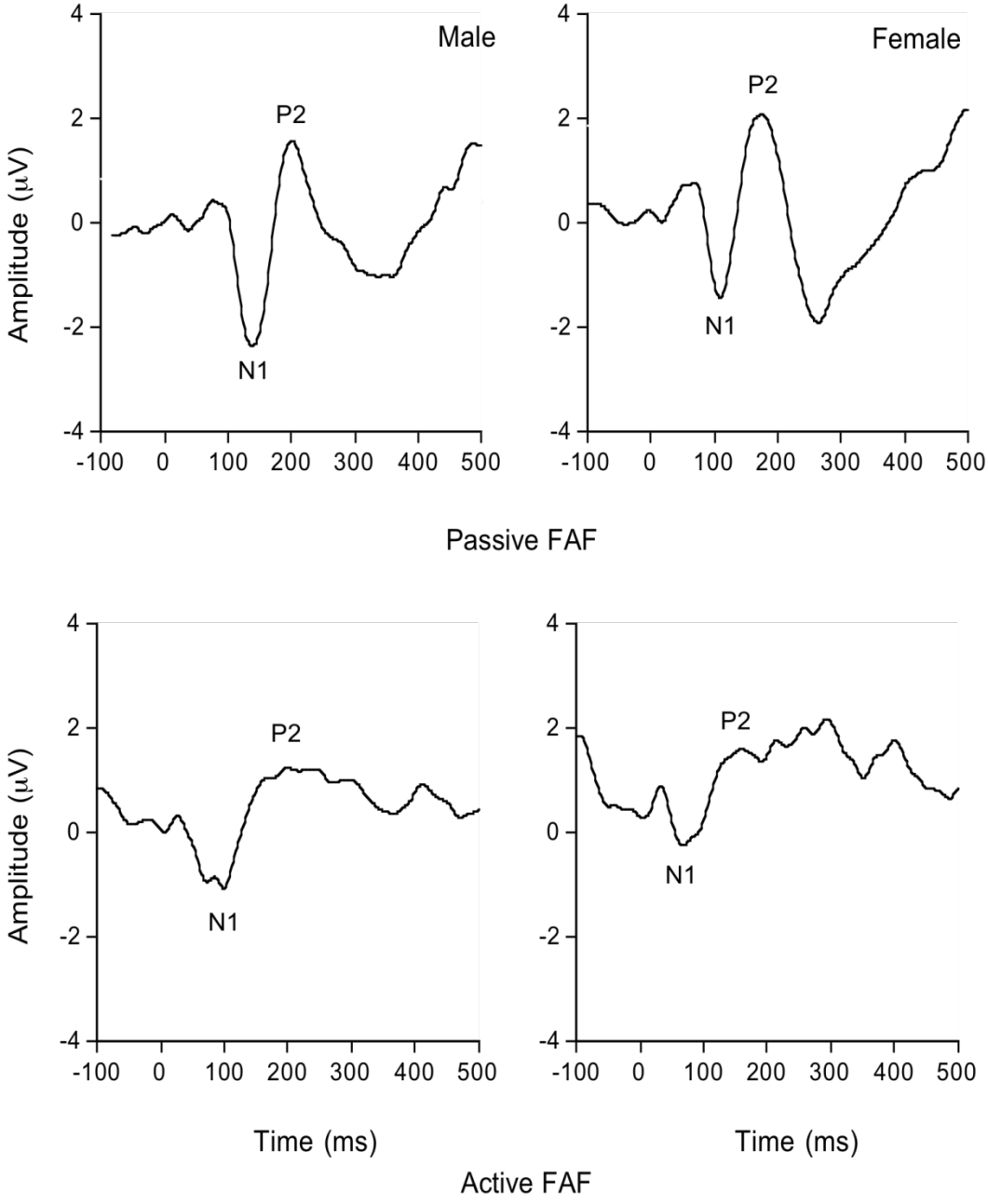


Figure 84. Grand average waveforms for the passive FAF condition compared to the grand average waveforms for the active FAF condition.

(i.e., NAF and FAF) conditions were statistically different ($p < .0001$). There were no significant differences between the passive NAF or FAF conditions ($p = 0.85$; cf. 116.7 ms vs. 117.6 ms).

P2 latency. P2 latency means and standard deviations are presented in Table 25 as a function of condition and gender and graphically presented in Figure 85. A three-factor mixed ANOVA was utilized to examine the effect of condition (i.e., active vs. passive), feedback (i.e., NAF vs. FAF), and gender on P2 latency (see Table 26). Significant main effects were found for condition, feedback, and gender. Two significant interactions were found, a two-way condition by feedback and a three-way interaction between condition by feedback by and gender. A three-way interaction plot for P2 latency is graphically presented in Figure 86.

Post hoc independent and pair-wise *t*-tests were undertaken to examine the source of the interaction. Independent *t*-tests examined differences across gender. A significant difference was found between the male and female passive NAF condition. These results indicated that during the passive NAF conditions P2 latencies were significantly shorter when recorded from female participants than male participants, $t(28) = 2.67$, $p = .012$. A statistically significant difference was also found between the male and female active FAF conditions. These results were interpreted to indicate that during the active FAF condition, P2 latencies were significantly shorter when recorded from female participants than male participants, $t(28) = 2.41$, $p = .023$. No significant differences were found between male and female P2 latencies collected during the passive FAF condition, $t(28) = 1.87$, $p = .072$ or the active NAF condition, $t(28) = 0.26$, $p = .80$.

Paired *t*-tests were conducted to evaluate P2 latency differences between matched conditions separately within each gender. For male participants, a significant P2 latency difference was found between the active NAF condition compared to the active FAF condition, $t(14) = 3.20$, $p = .006$). This result indicates that when recorded from male participants P2

Table 25.

P2 Mean Peak Latencies and Standard Deviations as a Function of Condition (i.e., Passive vs. Active), Feedback (i.e., NAF vs. FAF), and Gender (i.e., Male vs. Female).

		Latency (ms)	
Condition	Feedback	Male	Female
Passive	NAF	123.7	109.7
		(16.3)	(13.9)
	FAF	122.1	113.1
		(15.5)	(17.2)
Active	NAF	82.7	66.8
		(23.0)	(23.2)
	FAF	95.2	85.9
		(20.8)	(25.4)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

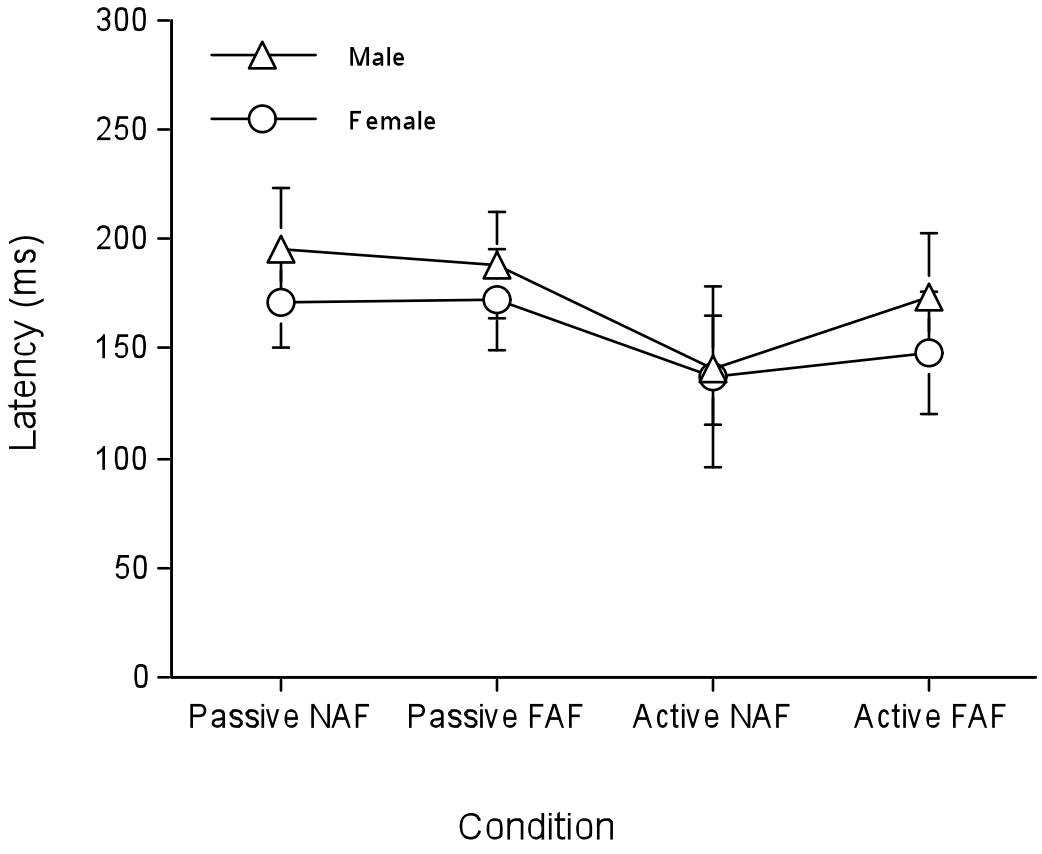


Figure 85. P2 mean peak latencies as a function of stimulus condition, feedback, and gender. Error bars represent +/- one *SD* of the mean.

Table 26.

Summary of Three-Factor mixed ANOVA Comparing Differences Between P2 Mean Peak Latency (ms) as a Function of Condition (i.e., Passive vs. Active), Feedback (i.e., NAF and FAF), and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	30337.20	1	30337.20	43.87	<.0001*	.61
Feedback	2726.53	1	2726.53	4.47	.043*	.14
Gender	8875.20	1	8875.20	5.83	.023*	.17
Condition X Gender	270.00	1	270.00	.39	.54	.01
Feedback X Gender	374.53	1	374.53	.62	.44	.02
Condition X Feedback	4416.53	1	4416.53	12.92	.001*	.32
Condition X Feedback X Gender	1672.53	1	1672.53	4.89	.035*	.15

Note. * significant at $p < .05$.

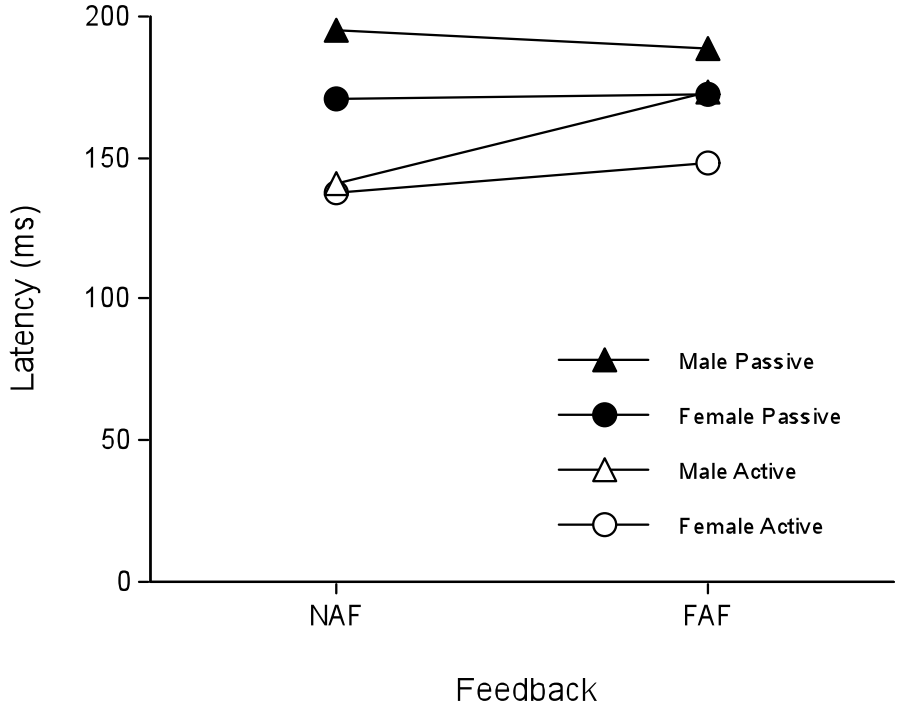


Figure 86. Interaction plot of P2 mean peak latency as a function of condition (i.e., active vs. passive), feedback (i.e., NAF vs. FAF), and gender.

latencies were significantly shorter for the active NAF condition compared to the active FAF condition. No significant P2 latency differences were found between the passive NAF and FAF condition, $t(14) = 1.21$, $p = 0.25$. For the female participants, no significant differences were found between the passive NAF condition compared to the FAF condition $t(14) = 0.24$, $p = .82$ or the active NAF condition compared to the FAF condition $t(14) = 1.14$, $p = 0.27$.

N1-P2 amplitude. Mean peak amplitudes and standard deviations, as a function of stimulus condition and gender, for the N1-P2 waveform complex are presented in Table 27. A three-factor mixed ANOVA was used to examine the effects of stimulus condition, feedback, and gender on N1-P2 amplitudes. The ANOVA results are summarized in Table 28. Only a significant main effect for condition was seen (see Figure 87). As evident in Figure 87, N1-P2 amplitudes were larger when recorded under the passive NAF and FAF conditions compared to the active NAF and FAF conditions.

Delayed Altered Auditory Feedback

Latency. Additional analysis was performed on the responses collected under the DAF 200 ms condition. During the DAF 200 ms condition, responses similar to those recorded under the other active conditions were observed (i.e., N1 and P2 waveform components); however, following these components an additional positive-negative-positive complex was present in all individual waveforms (see Figure 77 and 78). For simplicity, these waveform components are referred to as P1', N1' and P2'. Mean latencies and standard deviations for N1, P2, P1', N1', and P2' waveform components, as a function of gender are presented in Table 29. As evident in Table 29, the latencies of N1' and P2' waveform components were approximately 200 ms later than the N1 and P2 waveform components. The difference between N1 ($M = 76.13$, $SD = 15.72$)

Table 27.

N1-P2 Mean Amplitudes and Standard Deviations as a Function of Condition (i.e., Passive vs. Active), Feedback (i.e., NAF vs. FAF), and Gender (i.e., Male vs. Female).

		Amplitude (μV)	
Condition	Feedback	Male	Female
Passive	NAF	3.1	3.8
		(1.7)	(2.1)
	FAF	3.6	3.8
		(1.6)	(2.0)
Active	NAF	2.6	3.9
		(2.3)	(2.4)
	FAF	3.3	3.4
		(2.0)	(1.6)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

Table 28.

Summary of Three-Factor mixed ANOVA Comparing Differences Between N1-P2 Amplitude as a Function of Condition (i.e., Active vs. Passive), Feedback (i.e., NAF vs. FAF), and Gender (i.e., Male vs. Female).

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	144.34	1	144.34	17.14	<.0001*	.38
Feedback	.182	1	.182	.076	.78	.003
Gender	19.96	1	19.96	1.905	.178	.064
Condition X Gender	.17	1	.17	.020	.89	.001
Feedback X Gender	8.97	1	8.97	3.76	.063	.118
Condition X Feedback	1.32	1	1.32	.818	.373	.028
Condition X Feedback X Gender	.097	1	.097	.060	.81	.002

Note. * significant at $p < .05$.

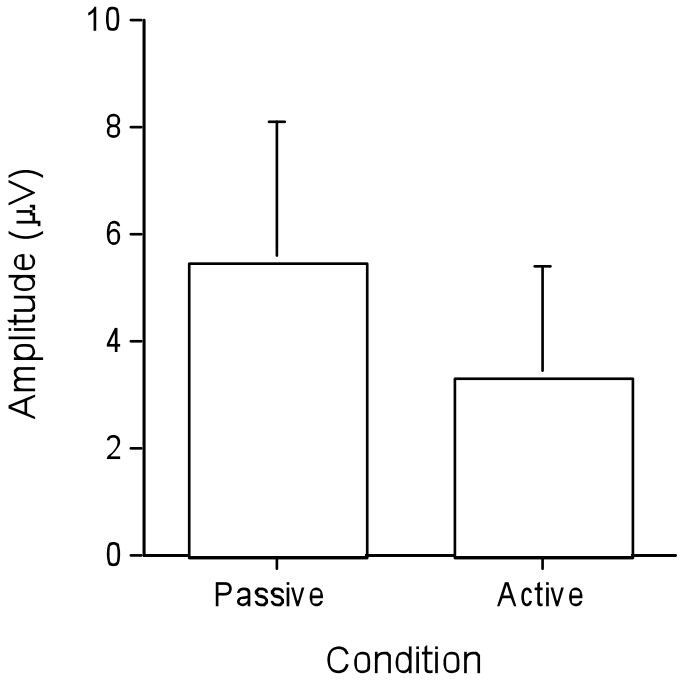


Figure 87. N1-P2 mean amplitudes as a function of condition (i.e., Passive vs. Active) collapsed across feedback and gender. Error bars represent + one *SD* of the mean.

Table 29.

Mean Peak Latencies and Standard Deviations as a Function Waveform Component (i.e., DAF 200 ms N1, P1, P1', N1' and P2') and Gender.

	Latency (ms)	
Waveform Component	Male	Female
N1	76.1	79.9
	(15.7)	(28.4)
P2	136.9	131.7
	(29.2)	(16.7)
P1'	231.2	252.7
	(21.3)	(22.2)
N1'	276.7	284.3
	(32.2)	(27.4)
P2'	344.8	340.3
	(51.5)	(22.3)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

and N1' (M = 276.67, SD = 32.19) for male participants is 200.5 ms and for female participants the N1 (M = 79.87, SD = 28.41) and N1' (M = 276.67, SD = 32.19) difference is 204.4 ms.

The difference between P2 (M = 136.93, SD = 29.19) and P2' (M = 344.80, SD = 51.49) for male participants is 207.9 ms and for female participants the P2 (M = 131.73, SD = 16.70) and P2' (M = 340.27, SD = 22.29) difference was 208.5 ms. The P1 component was absent in the majority of waveforms recorded in the active speaking conditions; therefore, only the N1 and P2 components were compared with the N1' and P2' components. The P1' component was present in all waveforms. Separate one-way ANOVAs were utilized to determine if there were significant differences between P1'-N1'-P2' latencies as a function of gender and between the N1 and P2 components subtracted from N1' and P2' components (i.e., latency difference) as a function of gender. ANOVA summaries are presented in Table 30. A statistically significant difference between male and female DAF 200 ms P1' responses was found (see Table 30).

Passive Self-produced Components Versus DAF 200 ms Components

Latency. Additional comparisons were made between the DAF 200 ms P1'-N1'-P2' components and the passive self-produced P1-N1-P2 components to determine if there were significant differences between component latencies. The absolute mean peak latencies for these two conditions are presented in Table 31. The difference between P1' and the self-produced passive P1 for the male group was 154.0 ms and for the female group was 183.2 ms. The difference between N1' and self-produced N1 for the male group was 152. 9 ms and for the female group was 174.5 ms. The difference between P2' and self-produced P2 for males was 149.9 ms and for females 169.5 ms. Separate one-way ANOVAs were undertaken to determine if there were significant differences between the passive and active component latencies as a

function of gender. ANOVA summaries are presented in Table 32. Only a significant difference between P1' and self-produced P1 male and female latencies were found (see Table 32).

Amplitude. Mean amplitudes and standard deviations for the passive self-produced /a/ P1-N1 and N1-P2 and the DAF 200 ms P1'-N1' and N1'-P2 complexes are presented in Table 33 as a function of gender. Two-factor ANOVAs were utilized to examine the differences between condition amplitudes as a function of gender. ANOVA summaries are presented in Tables 34 and 35. A significant main effect of condition was found for DAF 200 ms P1'-N1' amplitude compared to the self-produced P1-N1 amplitude and the DAF 200 ms N1'-P2' amplitude compared to the self-produced N1-P2 amplitude. That is, amplitudes recorded during the self-produced passive listening condition (i.e., P1-N1 and N1-P2) were larger than the "echo" response amplitudes (i.e., P1'-N1' and N1'-P2') recorded during the active DAF 200 ms condition (see Figure 88).

Discussion

In this experiment, electrophysiological measures were used to investigate the theoretical assumption that cortical suppressions during self-produced vocalizations are the result of neuronal matching between the actual feedback and the sensory expectations or corollary discharges. It was predicted that the responses recorded during passive listening conditions would differ from the responses recorded during conditions of active vocalizations. Specifically, passively evoked response amplitudes will be larger than the active conditions (Beal, Cheyne, Gracco, Quraan, Taylor, & De Nil, 2010; Curio et al., 2002; Ford, Mathalon, Heinks et al., 2001; Heinks-Maldonado et al. 2005; Houde et al., 2002; Kudo et al., 2004; Numminen et al., 1994; Ventura, Nagarajan, & Houde, 2009). As expected, N1-P2 amplitudes were reduced when recorded during active vocalization compared to responses recorded passively.

Table 30.

Summary of One-way ANOVA Comparing Latency Differences Between P1', N1', P2' and N1'-N1 and P2'-P2 as a Function of Gender.

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
P1'	3456.133	1	3456.133	7.29	<.012*	.21
N1'	433.20	1	433.20	.48	.49	.017
P2'	154.13	1	154.13	.098	.76	.003
N1'-N1	112.13	1	112.13	.162	.69	.006
P2'-P2	3.33	1	3.33	.002	.96	.000

Note. * significant at $p < .05$.

Table 31.

Mean Peak Latencies and Standard Deviations as a Function Waveform Component (i.e., Self-produced P1, N1, P2 and DAF 200 ms P1', N1' and P2') and Gender.

		Latency (ms)	
	Waveform Component	Male	Female
Passive	P1	77.2	69.5
		(13.6)	(18.4)
	N1	123.7	109.7
		(16.3)	(13.9)
	P2	194.9	170.8
		(28.7)	(20.1)
Active	P1'	231.2	252.7
		(21.3)	(22.2)
	N1'	276.7	284.3
		(32.2)	(27.4)
	P2'	344.8	340.3
		(51.5)	(22.3)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

Table 32.

Summary of One-way ANOVA Comparing Latency Differences Between DAF 200 ms and Passive Self-Produced Listening Components (i.e., P1'- P1, N1'-N1, and P2'-P2) as a Function of Gender.

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
P1'- P1	6394.80	1	6394.80	7.29	<.009*	.22
N1'- N1	3499.20	1	3499.20	2.96	.096	.096
P2'- P2	2881.20	1	2881.20	1.35	.26	.046

Note. * significant at $p < .05$.

Table 33.

Mean Amplitudes and Standard Deviations of the Passive Self-Produced Condition (i.e., P1-N1 and N1-P2) and the DAF 200 ms Condition (i.e., P1'-N1' and N1'-P2') as a Function of Gender.

	Latency (ms)	
Stimuli	Male	Female
P1-N1	3.1	3.8
	(1.7)	(2.1)
N1-P2	5.0	6.3
	(2.5)	(2.6)
P1'-N1'	1.31	1.4
	(0.9)	(1.0)
N1'- P2'	2.5	2.8
	(1.5)	(1.6)

Note. Values enclosed in parenthesis represent one standard deviation of the mean.

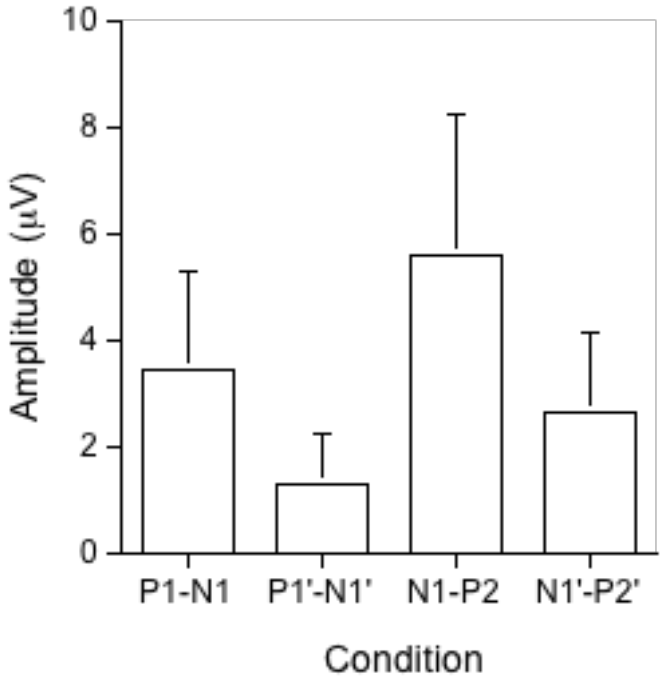


Figure 88. Mean amplitude (μV) as a function of condition, passive self-produced (P1-N1 and N1-P2) compared to active DAF 200 ms (P1'-N1' and N1'-P2'), collapsed across gender. Error bars represent + one *SD* of the mean.

Table 34.

Summary of Two-Factor Mixed ANOVA Comparing Differences Between DAF 200 ms and Passive Self-produced Listening Component Amplitudes (i.e., P1'-N1' vs. P1-N1) as a Function of Gender.

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	66.89	1	66.89	33.90	<.0001*	.55
Gender	1.69	1	1.69	.694	.41	.024
Condition X Gender	1.19	1	1.19	.605	.44	.021

Note: * significant at $p < .05$

Table 35.

Summary of Two-Factor Mixed ANOVA Comparing Differences Between DAF 200 ms and Passive Self-produced Listening Component Amplitudes (i.e., N1'-P2' vs. N1-P2) as a Function of Gender.

Source	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Condition	131.56	1	131.56	43.96	<.0001*	.61
Gender	10.26	1	10.26	1.69	.20	.057
Condition X Gender	4.59	1	4.59	1.53	.23	.052

Note: * significant at $p < .05$

With respect to component latency it was predicted that there would not be any significant latency differences found for these experimental conditions. This hypothesis was generated given that the paradigm employed by Heinks-Maldonado et al., (2005) was modified and adapted for this experiment. In that study, Heinks-Maldonado et al. (2005) did not find significant latency differences.

Contradictory to the proposed hypothesis, significant latency differences were found for both N1 and P2 components. Other researchers have reported equivocal findings, with respect to latency differences recorded via passive listening and vocalization. Curio et al. (2000) showed a left hemisphere specific decrease in M100 latencies recorded via passive listening compared to self-vocalizations. Also, Gunji, Hoshiyama, and Kakigi (2000) recorded vocalization related cortical potentials and N1 and P2 components simultaneously and found that latencies were only significantly different after vocalization began (i.e., "post vocalization").

Also, it has been speculated that auditory neural suppression varies depending on the incoming auditory information (Aliu et al., 2008). For example, signals that are more closely related to the expected auditory feedback are suppressed (Heinks-Maldonado et al., 2005; Houde et al., 2002), possibly, as a way of distinguishing between self-vocalizations and exogenous auditory input**.** Several of the results found here were unexpected and in general, experimental findings contradict most of the initially formulated hypotheses. Namely, in this experiment there were no significant differences in N1-P2 amplitude between the feedback conditions. Based on the results of previous electrophysiological studies (Ford et al., 2002, Heinks-Maldonado et al., 2005; Houde et al., 2002) it was hypothesized that significantly smaller amplitudes would be recorded when the participant was presented with NAF (i.e., feedback was not disrupted) compared to the FAF. Further, based on behavioral findings that longer DAF conditions (i.e., >

50 ms) produced “stuttering like dysfluency” in normally fluent speakers (Stuart et al., 2002) it was posited that the amplitudes recorded via the NAF condition would be significantly smaller than component amplitudes recorded via the DAF 200 ms altered conditions. Moreover, it was speculated that FAF and DAF 200 ms conditions would be similar in terms of component latency and amplitude.

There are several plausible explanations for the lack of N1-P2 amplitude differences found in this experiment compared to previous works (Curio et al., 2000; Heinks-Maldonado et al., 2005; 2007; Houde et al., 2002). Theoretically, feedback controls act to tune the feedforward control subsystem when errors are produced (Guenther, 2001; Guenther & Ghosh, 2001). Such that, the corrective motor commands are assimilated in to the feedforward commands and utilized for future executions of the motor act (Guenther, 2006, 2007). Once, the feedforward commands are tuned (i.e., motor adjustments learned and incorporated) the reliance on feedback commands are reduced and feedforward control subsystems function to facilitate actions. Given that auditory feedback was not randomized in terms of individual presentations, it is possible that the participant became rapidly accustom to the feedback perturbation, which in turned was assimilated into the feedforward commands, and resulted in proposed corollary discharge matching. In other words, the feedback perturbation became the expected sensory input, which matched the AAF condition.

Predictably the sensory input has also been shown to effect cortical responsiveness. In this experiment, participants triggered the feedback presentations when they spoke into the microphone. Schafer and Marcus (1973) reported that responses recorded while the participant self-produced tones had smaller amplitudes and shorter latencies than those recorded while passively listening to tones. However, in a subsequent study these authors related amplitude

differences to the predictability of the self-generated stimulus (Schafer, Amochaev, & Russell, 1981). Therefore, the feedback event was predictable and ultimately governed by the participants.

The predictability of feedback may have also consequently affected listener attention. Although, N1 and P2 components are classified as exogenous responses that represent stimulus encoding without active listener attention (Martin et al., 1997; Ostroff et al., 1998; Sharma & Dorman, 1999, 2000; Whiting et al., 1998), numerous studies have shown that these components are, in fact, affected by attention. N1 amplitude increases with listener attention (Davis, 1964; Hillyard et al., 1973; Keating & Ruhm, 1971; Näätänen et al., 1978; Picton & Hillyard, 1974; Roth et al., 1976) and P2 decreases with attention (Michie et al., 1993; Näätänen & Picton, 1987). Therefore, participant attention may have waned throughout the recording, which adversely reduced component amplitude.

Heinks-Maldonado et al. (2005) noted the possible effect of task predictability and attention. During vocalization tasks, the participant was presented with a random set of feedback conditions, consisting of NAF, frequency shifted tokens, or a token produced by another listener. These authors suggested that, although unlikely, it is possible that when participants heard an unexpected feedback (i.e., not there own voice) they allocated more attention to those tasks, which could have led to increased N1 amplitude. Whereas, during the unaltered task the participant spoke and heard their auditory feedback without an alteration (i.e., as expected), which might have contributed to the dampening of the N1 amplitudes.

Another unexpected experimental finding was the effect of DAF 50 ms on N1 and P2 latency. Given that Heinks-Maldonado et al. (2005) did not report significant N1 latency differences between altered feedback conditions, stable N1 and P2 latencies were also expected

across these experimental conditions. In this study, some latency differences were seen across the NAF, FAF, and DAF 200 ms conditions; however, for both males and females, N1 and P2 latencies were generally not significantly different. Thus, these findings are essentially consistent with those reported by Heinks-Maldonado et al., (2005).

Conversely, when recorded from male participants N1 and P2 latencies in the DAF 50 ms were significantly longer than the other three feedback conditions. This consistent prolongation of component latency in the DAF 50 ms condition was not seen in responses recorded from the female participants. Given, that one of the purposes of this experiment was to investigate wave component gender differences, these findings were of interest. However, devising a sound explanation as to why male participant latencies would be more susceptible to the DAF 50 ms alteration than female participants is difficult; especially, considering the lack of research investigating the underlying cortical processing mediating the effects of DAF. Although, there have been several brain imaging investigations employing DAF (Hashimoto & Saki, 2003; Hirano, et al., 1996; 1997; Takaso et al., 2010), to the best of this author's knowledge, this is the only study that has used DAF to record auditory evoked potentials during self-produced vocalizations.

One could reason that males and females process temporal alternations differently. Notably, behavioral studies have shown that the fluent speech of males is more adversely affected than the fluent speech of females (Bachrach, 1964; Fukawa et al., 1988; Sutton et al., 1963). Corey and Cuddapah (2008) speculated that the increased DAF effects in males might be correlated with the increased prevalence of developmental stuttering in males versus females. Further, several observations of these results can be made: 1) responses recorded using DAF 50 ms were not similar to those recorded using NAF and 2) responses recorded using DAF 200 ms

were not similar to those recorded using FAF. Therefore, both of these findings contradict the initial experimental assumptions.

Moreover, the argument could be made that the latency differences found throughout this investigation are related to timing differences introduced as a result of the self-vocalized triggering method used herein. Houde et al. (2002) used this rationale to account for latency differences between their experimental results and those found in Curio et al. (2000). They posited that a trigger that is activated by exceeding a voltage threshold would also be susceptible, in terms of the "triggering time", to amplitude modulations from the speaker. If the speaker changed the loudness of their voice, the trigger would come on at a faster rate (i.e., the voltage threshold would be reached more quickly) than for softer inputs (i.e., there would be a lag in the time it took to reach the voltage threshold). However, Curio et al. (2000) and Houde et al., (2002) showed shorter latencies for the listening conditions compared to the speaking conditions. In this experiment the opposite was shown; latencies for the speaking conditions were shorter than the listening conditions and NAF latencies were shorter than the latencies in the AAF conditions. According to the rationale proposed by Houde et al., participant vocalizations may have been softer when presented with NAF and louder for AAF conditions.

The listening and speaking conditions were presented using different triggering methods; therefore, the effect of intensity, ISI, and token durations must be considered. These parameters were not as controlled during the active speaking tasks as during the passive listening tasks. During the passive listening tasks the variables were consistent between each presentation. However, during the speaking tasks, recordings were triggered at the onset of participant voicing via a mechanical trigger, thus, the participant controlled the ISI between stimulus presentations, the duration of the actual token, and the intensity of the token. Although, prior to testing the

participant was trained to vocalize at the desired rate but intensity inter- and intra-participant variability was inevitable.

Additionally, in must be noted that responses recorded to self vocalizations may be influenced by changes in vocal pitch, rate, and intensity (Munhall, 2001) as a way of compensating for the feedback perturbations (Tourville et al., 2008) and/or bone conducted signal presented to the speaker upon vocalization (Mayer et al., 2000). In this light, it is possible that these factors increased and decreased response latencies accordingly.

Delayed Auditory Feedback 200 ms

During the DAF 200 ms condition an additional positive-negative-positive complex was recorded approximately 200 ms following the N1 and P2 components and are referred to here as P1', N1' and P2'. Overall, no significant differences were seen in terms of component latency between male and female participants, with the exception of a significant P1' latency gender difference. It was speculated that these additional components might be the result of hearing your own voice with enough of a delay that it emulates passively presented auditory information (i.e., an echo). Therefore, P1', N1', and P2' were compared to the passive self-produced P1, N1, and P2 components. The only significant difference was between male and female P1' latencies and the self-produced passively recorded P1 latencies. With respect to amplitude there was a significant main effect of condition between both P1-N1 amplitudes compared to P1'-N1' amplitudes and N1-P2 amplitudes compared to N1'-P2' amplitudes. Given these results it is possible that these components are in fact the result of the additional acoustic information presented via the delayed feedback.

CHAPTER IV: GENERAL DISCUSSION

Summary of Experimental Findings

In the first experiment, it was hypothesized that P1-N1-P2 component latencies and amplitudes would differ when recorded with nonspeech stimuli compared to speech stimuli and that no significant differences would be found between the passively presented speech tokens. Using a non-speech tone burst and various natural and synthetic vowel tokens it was shown, that P1-N1-P2 components could be reliably and consistently recorded with each stimulus without gross morphological variations. It was also found that component latencies were significantly shorter when recorded with natural tokens versus synthetic tokens and with tones versus speech tokens. For the natural versus synthetic tokens, it was concluded that, since temporal characteristics were generally the same, these differences are related to differences between the spectral characteristics of the natural tokens versus the synthetic tokens. It was also concluded that for the tone versus speech conditions amplitude and latency differences were related to both temporal and spectral aspects of the token.

In the second experiment, it was shown that responses recorded during self-vocalizations were reduced in comparison to those recorded passively. This finding supports the notion of feedforward and feedback monitoring of self-produced speech. Theoretically, incoming auditory inputs are monitored via feedforward and feedback control subsystems (Guenther, 1994, 1995, 2001, 2007), which allow for the detection of self-produced errors (Heinks-Maldonado et al., 2005) and the monitoring of incoming auditory information as a way of distinguishing between internal and external signals (Ford et al., 2002; Frith, 1992).

This model was also examined in terms of the selectively of auditory cortical responses. If the proposed theoretical model held true (Guenther, 2006; Heinks-Maldonado et al., 2007),

NAF heard by participants during active vocalizations, should engage vocalization-induced suppression of the auditory evoked potentials demonstrated by a reduction in component amplitude (Aliu et al., 2008). The contrary (i.e., unsuppressed component amplitudes) should be seen when sufficient, real time feedback perturbations are presented (e.g., FAF and DAF). When the expected feedback does not match the actual feedback, sensory to motor suppression will not occur, and there will not be a reduction in amplitude. Further, it has been proposed that more cortical suppression will be induced the closer the actual feedback is to the expected feedback. This was not seen in this experiment. All conditions recorded with self-vocalized tokens were suppressed and there were not significant amplitude differences between these conditions. Nevertheless, these findings do not necessarily refute the proposed model, given that the lack of amplitude differences can be attributed to differences across studies.

Responses recorded under DAF 200 ms condition were interesting. As with the other conditions, a consistent albeit reduced N1-P2 complex was seen; however, an additional positive-negative-positive complex (i.e., termed P1', N1', and P2') was seen at approximately 200 ms after the initial components. It was founded that P1', N1', and P2' component characteristics were similar to those recorded during the passive self-produced condition; therefore, it was concluded that the 200 ms was such that the auditory input emulated passively incoming auditory signals (i.e., an echo). A shorter DAF 50 ms condition was also employed. Based on behavioral studies, showing that this delay does not cause the disruptions in fluent speech seen with longer delays (i.e., DAF 200 ms; Stuart et al., 2002) it was hypothesized that component amplitudes and latencies would be similar to those recorded with the NAF condition. Again, this was not found in the experimental data collected here. In considering the discrepancy between behavioral findings and electrophysiological findings, for example, normally fluent

participants presented more stuttering like dysfluencies under longer feedback conditions; however, amplitude differences were not found for the wave components recorded here, one could reason that proprioceptive or somatosensory factors may play a more significant under DAF in individuals who are fluent.

Does Stimulus Condition Effect Auditory Monitoring?

The human auditory system is responsible for the effective simultaneous processing of a vast amount of incoming auditory inputs that allow for the monitoring of both self-produced and externally produced signals. If auditory inputs could not be distinguished as self-produced or externally generated, both humans and nonhumans would be at substantial evolutionary disadvantage, simply because approaching dangers would not be recognized as such. According to the proposed model, feedforward and feedback control subsystems work in concert a facilitate auditory monitoring. Additionally, it has been shown that cortical responses differ when elicited via speech and nonspeech stimuli (Aiken & Picton, 2008a, 2008b); however, the underlying cortical processing and the functional anatomy involved is not clearly understood at this time.

Researchers have hypothesized that the auditory system is innately tuned to encode the properties of speech (Liberman & Mattingly, 1989) and there is a "special centre" in the brain that recognizes speech sounds over nonspeech sounds. These researchers deem that "speech is special". This hypothesis has been questioned by those presenting the alternative explanation that spectral and temporal features of the signal drive processing, especially, at levels below the auditory cortex (Uppenkamp et al., 2006). It is unclear whether these speech sounds are recognized innately as such (i.e., is speech special?) or if the central auditory system analyzes speech and nonspeech sounds in terms of the spectro-temporal characteristics disregarding the

linguistic content of the signal, until higher cerebral areas. Additionally, it is unclear if synthetic speech or naturally produced speech is processed differently.

What is the answer to, *Are responses to nonspeech stimulus different from those recorded with speech tokens?* From the findings of Experiment 1, the answer would be yes. P1-N1-P2 latencies were shorter and P1-N1 and N1-P2 amplitudes larger when responses were elicited with the nonspeech token compared to the speech tokens. What is the answer to, *Are response to natural vowel tokens different from those recorded with synthetic vowel tokens?* Again, the answer would also be yes. It is highly probable differences found in this experiment were related to the spectro-temporal parameters of the tokens. Further research is warranted to investigate cortical speech processing. Specifically, more research is needed to parse out if speech is innately and initially recognized whereby the auditory stimuli dominates initial processing or if speech is processed in the same manner as other incoming auditory signals and recognized and encoded as speech at higher cortical levels. Further, additional research is needed to determine if synthetic speech is in fact processed the same as natural speech.

It also must be noted that responses recorded during the active speaking conditions (see Figures 63 to 78) were more noisy than those recorded during the passive listening conditions (see Figures 30 to 57). When recording ERPs it is important to ensure that the recorded responses represent synchronous neural events and are not confounded or degraded by the presence of extraneous noise (Glaser & Ruchkin, 1976; Perry, 1966). One could argue that N1-P2 amplitude differences between passively elicited responses and actively elicited responses are the result of degraded S/Ns and are not related to neural suppression. Several measures were taken to reduce the presence of noise during the active speaking conditions such as asking the participant to limit head movements and eye movements (see Appendix H), monitoring and

removing eye movement artifacts, and selecting the vowel /a/ which requires limited jaw movement for production. However, it is obvious when looking at the recorded waveforms that those elicited during active conditions are more noisy than those recorded during the passive conditions.

Do Males and Females Process Auditory Information Differently?

It was a goal of this experiment to examine the differences in cortical responses between male and female participants. When recording early latency potentials (e.g. ABRs) responses recorded from females typically have shorter latencies and larger amplitudes than those recorded from males (Beagley et al., 1978; Berghotlz, 1981; Chan et al., 1988; Don et al., 1993). With respect to auditory long latency ERPs reports examining gender differences have been equivocal. Several investigators have reported gender differences (Altmann & Vaitulevich, 1990; Onishi & Davis, 1968), while others have reported no significant differences between males and females (Polich et al., 1985). Gölgeli et al. (1999) reported that N1-P2 and N2-P3 interpeak amplitudes were larger in males than females. While no significant gender differences were reported for N1, P2, N2, P3 amplitudes or latencies. Kudo et al. (2004) specifically evaluated gender differences when recorded during a selective attention task presented during passive listening and when performing the same task while overtly producing the vowel /a/. These authors did not find any significant gender differences for the N1 wave component. Data for the other waveform components was not reported.

In the current dissertation, gender differences were more pronounced in the active speaking conditions compared to the passive listening conditions. In the passive conditions, significantly shorter P2 latencies were found in female participants compared to male participants. This finding is consistent with the trend seen for ABRs. Researchers have

speculated that gender differences are the result of the anatomical variations between males and females (Don et al., 1993). For example, it has been indicated that males have larger heads and thicker skulls than females (Trune et al., 1988; Stockard, Stockard, & Sharborough, 1978) and females have shorter auditory neural pathways (Stockard et al., 1978; Stockard, Hughes, & Sharborough, 1979) and structurally smaller cochlea than males (Don et al., 1993; Sato et al., 1991), which in turn, respectively, results in the reduced amplitudes in males and the shorter latencies in females.

However, if differences between male-female anatomies were solely responsible for the shorter P2 latencies found here, a main effect for gender would have been seen across all components and not the P2 selectively. Given this, it is possible that this finding reflects the overall difference in auditory processing speed between males and females. In the active conditions, the DAF 50 ms condition significantly increased response latencies, so that, latencies recorded under this condition were significantly longer than the other conditions. This trend was not seen in female participants. Thus, it could be speculated that auditory processing differs between males and females, upon engaging the speech mechanism.

Future Research Directions

The current findings support the speculation that feedforward and feedback systems mediate auditory motoring through the motor-to-sensory matching. However, the need for further research is apparent. Research is needed to examine not only normal cortical processing of speech production and perception, but also cortical speech processing in those with speech motor and fluency disorders. Guenther and colleagues (Civier & Guenther, 2005; Max, Guenther, Gracco, Ghosh, & Wallace, 2004**)** have applied the DIVA model to the disorder of

stuttering; however, disorder of stuttering is still poorly understood. Future research will be directed toward the further investigation of the disorder of stuttering.

Stuttering is a speech disorder where the speaker knows what he would like to say but is unable to produce the words. Typically, IWS produce speech, which is interrupted by part-word repetitions, prolongations, and/or audible fixations (Dayalu, Kalinowski, & Stuart, 2005). To date, there is no known cause or cure for speech and results from investigations focusing on delineating the cause of stuttering are equivocal. It has been evident for more than 50 years that stuttering is significantly reduced when IWS speak under conditions of AAF (Lee, 1950). These conditions include DAF (Naylor, 1953; Kalinowski et al., 1993; Kalinowski et al., 1996), FAF (Howell et al., 1987; Kalinowski, et al., 1993; Hargrave et al., 1994; Stuart et al., 1996; Stuart, Kalinowski, & Rastatter, 1997), masked auditory feedback (Maraist & Hutton 1957; Kalinowski et al., 1993), and reverberation (Adamczyk, Sadowska, & Kuniszyk-Jozkowiak, 1975). However, the method in which AAF reduces stuttering is undetermined.

Numerous researchers have investigated both the peripheral and central auditory system involvement. Initially, it was suspected that aberrant peripheral auditory structures caused a disruption in speech-auditory monitoring (Cherry & Sayers, 1956; Stromsta, 1972; Webster & Dorman, 1970). However, a defendant peripheral abnormality has yet to be found. Several central auditory system theories have also been proposed. Fairbanks and Guttman (1958) hypothesized that stuttering results from a deficit in the rapid processing of auditory feedback.

Recently, imaging studies have confirmed that the auditory system is crucial for fluent speech production through an auditory feedback mechanism (Hirano et al., 1997; Ford & Mathalon, 2004; Heinks-Maldonado et al., 2005, 2006). These researchers have suggested that forward flowing speech is the result of a matching between the corollary discharge produced

when a speech motor command is executed and the processing of auditory feedback received during self-produced speech (Heinks-Maldonado et. al., 2005). This corollary discharge is the prediction of self-produced sensory input as a result of the motor command. In other words, an individual formulates the motor command for speech in the frontal lobe. During production, a speech motor command is executed by the sensorimotor system; an "efferent copy" is created which produces corollary discharge representing expected sensory input. When the speech command is executed auditory feedback of the spoken message is heard at the level of the peripheral auditory system and then processed through the central auditory pathway to the auditory temporal lobe. A comparison is then made between the spoken message (i.e., auditory feedback) and the corollary discharge (i.e., predicted input). If a match occurs between the two signals (i.e., there is no discrepancy), forward flowing speech will occur (Ford & Mathalon, 2004; Heinks-Maldonado et al., 2005, 2006) and auditory cortical activity is reduced signaling the auditory feedback is self-generated. It is hypothesized that disruptions in the matching of corollary discharge to actual received auditory input plays a role in the production of stuttered speech production. In that, aberrant sensory cancellation occurs when stuttered speech is produced.

To investigate this hypothesis, the general focus of future proposed research would be to investigate the relationship between auditory function in individuals with persistent developmental stuttering and IWF. Two overall research questions will be asked: Are auditory electrophysiological responses different in individuals with persistent developmental stuttering and IWF? Does condition effect auditory electrophysiological responses (i.e., will responses differ if measurements are recorded under listening or speaking conditions)? To address these proposed research questions a series of experiments will be performed. It is hypothesized that

anatomical structures and physiological processes are similar between IWS and IWF when the speech mechanism is not engaged. Thus, it is predicted that similar electrophysiological responses will be recorded from IWS and IWF.

The question of whether or not electrophysiological responses recorded from IWS and IWF differ when the speech mechanism is not engaged, will also be addressed in future research. Specifically, the question will be asked "Are auditory electrophysiological P1-N1-P2 responses different in IWS and IWF when evoked using recorded speech stimuli and the participant is silent?" It is again hypothesized that anatomical structures and physiological processes are similar between IWS and IWF when the speech mechanism is not engaged. Based on this assumption, it is predicted that similar P1-N1-P2 responses will be elicited from IWS and IWF. Therefore, implementation of this experiment will further establish similarities between IWS and IWF in the absence of speech production.

Another possible question is "Will auditory evoked responses recorded from IWS and IWF differ when evoked using self-produced non-altered feedback?" It is hypothesized that response differences are due to the mismatching of actual auditory input and expected sensory input. If the proposed hypothesis holds true measurement under this condition should result in a dampening of P1-N1-P2 activity in the control group (i.e., IWF) representing auditory cortical inhibition to self-produced speech. However, in the experimental group (i.e., individuals who stutter) this pattern of auditory cortical inhibition should be absent due to the fact that stuttered speech does not match the expected sensory input (i.e., corollary discharge).

Further experimental conditions presenting altered auditory feedback during active vocalizations should be implemented. It is predicted that group differences will be recorded. Altered auditory feedback has shown to have adverse effects on the production of fluent speech

(Lee, 1950). In that, when altered auditory feedback is presented to a fluent speaker stuttering like behaviors will be produced. In IWS altered auditory has shown to immediately inhibit stuttering so that fluent speech can be produced. This inhibitory effect, thus, provides evidence for a relationship between the auditory system and speech production.

REFERENCES

Ackermann, H., & Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language, 89,* 320-328.

Adamczyk, B., Sadowska, E., & Kuniszyk-Jozkowiak, W. (1975). Influence of reverberation on stuttering. *Folia Phoniatrica*, *27,* 1-6.

Adler, G., & Adler, J. (1989). Influence of stimulus intensity of AER components in the 80- to 200-millisecond latency range, *Audiology, 28,* 316-324.

Adrian, E. D., & Matthews, B. H. C. (1934). The Berger rhythm: Potential changes from the occipital lobes in man. *Brain, 57,* 355-385.

Agung, K., Purdy, S.C., McMahon, C.M., & Newall, P. (2006). The use of cortical auditory evoked potentials to evaluate neural encoding of speech sounds in adults. *Journal of the American Academy of Audiology, 17,* 559-572.

Aiken, S. J., & Picton, T. W. (2008a). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research, 245,* 34-47.

Aiken, S. J., & Picton, T. W. (2008b). Human cortical responses to the speech envelope. *Ear and Hearing, 29,* 139-157.

Alain, C., & Tremblay, K. (2007). The role of event-related brain potentials in assessing central auditory processing. *Journal of the American Academy of Audiology, 18,* 573-589.

Alain, C., Woods, D. L., & Covarrubias, D. (1997). Activation of duration-sensitive auditory cortical fields in humans. *Electroencephalography and Clinical Neurophysiology, 104,* 531-539.

Alario, F. –X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area in word production. *Cognitive Brain Research, 1076,* 129-143.

Alexander . J. E., & Polich, J. (1997). Handedness and P300 from auditory stimuli. *Brain and Cognition,* 35, 259-270.

Alho, K., Sams, M., Paavilainen, P., & Näätänen, R. (1986). Small pitch separation and the selective-attention effect on the ERP. *Psychophysiology, 23,* 189-197.

Aliu, S. O., Houde, J. F., &Nagarajan, S. S. (2008). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience, 21,* 791-802.

Altmann, J. A., & Vaitulevich, S. F. (1990). Auditory image movement potentials. *Electroencephalography and Clinical Neurophysiology, 75,* 323-333.

American National Standards Institute. (2004). *Methods for manual pure-tone threshold audiometry* (ANSI S3.21-2004). New York: ANSI.

American National Standards Institute. (1999). *Permissible ambient noise levels for audiometric test rooms.* (ANSI S3.1-1999). New York: ANSI.

American Speech-Language-Hearing Association. (1988). Guidelines for determining threshold level for speech. *ASHA, 30,* 85-89.

American Speech-Language-Hearing Association. (1996). Central auditory processing: Current research and implications for clinical practice. *American Journal of Audiology, 5,* 41-54.

Antinoro, F., Skinner, P.H., & Jones, J.J. (1969). Relation between sound intensity and amplitude of the AER at different stimulus frequencies. *Journal of the Acoustical Society of America, 46,* 1433-1436.

Assmann, P.F., & Katz, W.F. (2000). Time-varying spectral change in the vowels of children and adults. *The Journal of Acoustical Society of America, 108,* 1856-1866.

Bachrach, D. L. (1964). Sex differences in reactions to delayed auditory feedback. *Perceptual and Motor Skills, 19*, 81-82.

Ballantyne, D. (1990). *Handbook of audiological techniques.* London: Butterworth-Heinemann.

Baran, J. A., Verkest, S., Gollegly, K., Kibbe-Michal, K., Rintelmann, W. F., & Musiek, F. F. (1985). Use of compressed speech in the assessment of central nervous system disorder. *Journal of the Acoustical Society of America, 78* (Suppl. 1), S41.

Bates, J.F., & Goldman-Rakic, P.S. (1993). Prefrontal connections of medial motor areas in the rhesus monkey. *Journal of Comparative Neurology, 336,* 211-228.

Beagley, H. A., & Knight, J. J. (1967). Changes in auditory evoked response with intensity. *Journal of Laryngology & Otology, 81,* 861-873.

Beal, D.S., Cheyne, D.O., Gracco, V., Quraan, M.A., Taylor, M.J., De Nil, L.F. (2010). Auditory evoked fields to vocalization during passive listening and active generation in adults who stutter. *NeuroImage.* doi:10.1016/j.neuroimage.2010.04.277

Behroozmand, R., Karvelis, L., Liu, H., & Larson, C.R. (2009). Vocalization-induced enhancement of the auditory cortex responsiveness during voice F(0) feedback perturbations. *Clinical Neurophysiology, 129,* 1303-1312.

Beine, B. (2007). Neurophysiologic basis of sleep and wakefulness. In N. Butkov & T. L. Lee-Chiong (Eds.), *The fundamentals of sleep technology* (pp.11-17). Philadelphia, PA: Lippincott Williams and Wilkins.

Bellis, T.J. (2003). *The assessment and management of central auditory processing disorders in the educational setting: From science to practice* (2nd ed.). San Diego, CA: Singular Publishing Group, Inc.

Bentall, R. P. (1990). The illusion of reality: A review and integration of psychological research on hallucination. *Psychological Bulletin, 107,* 82-95.

Berger, K. W. (1978). Speech audiometry. In D. Rose (Ed.), *Audiological assessment* (2nd ed., pp. 227-260). Englewood Cliffs, NJ: Prentice-Hall.

Bergholtz, L. (1981). Normative data in clinical ABR. *Scandinavian Audiology, 13,* 185-190.

Berlin, C. L., Lowe-Bell, S. S., Cullen, Jr., J. K., Thompson, S.L., & Loovis, C.F. (1973). Dichotic speech perception: An interpretation of right-ear advantage and temporal offset effects. *Journal of the Acoustical Society of America, 53,* 699-709.

Bess, F. (1983). Clinical assessment of speech recognition. In D. Konkle & W. Rintelmann (Eds.), *Principles of speech audiometry* (pp. 127-201). Baltimore, MD: Academic.

Bhatnagar, S. C., & Andy, O. J. (1995). *Neuroscience for the study of communicative disorders.* Baltimore, MD: Williams & Wilkins.

Bias (2003). Peak Version 4.13 User Guide. California.

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition, 39,* 173-194.

Blackmore, S. J., Rees, G., & Firth, C. (1998). How do we predict the consequences of our actions? A functional imaging study. *Neuropsychologia, 36*, 521-529.

Boberg, E., Yeudall, L., Schopflocher, D., & Bo-Lessen, P. (1983). The effects of an intensive behavioral program on the distribution of EEG alpha power in stutterers during the processing of verbal and visuospatial information. *Journal of Fluency Disorders, 8,* 245-263.

Bocca, E., Calearo, C., & Cassinari, V. (1954). A new method for testing hearing in temporal lobe tumors. *Acta Otolaryngology, 44,* 219-221.

Boersma, P., & Weenink, D. (2008). Praat: Doing phonetics by computer. (Version 5.0.32) [Computer Program]. Retrieved August 12, 2008, from http://www.praat.org/

Borden, G. J. (1979). An interpretation of research on feedback interruption in speech *Brain and Language, 7*, 307-319.

Brenowitz, E. A., Margoliash, D., & Nordeen, K. W. (1997). An introduction to birdsong and avian song system. *Journal of Neurobiology, 33,* 495-500.

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology, 47,* 191-196.

Broadbent, D. E., & Gregory, M. (1964). Accuracy or recognition for speech presented to the right and left ears. *Quarterly Journal of Experimental Psychology, 16,* 359-360.

Brunt, M. (2002). Cochlear and retrocochlear behavioral tests. In J. Katz (Ed.), *Handbook of clinical audiology* (pp. 111-123). New York: Lippincott Williams & Wilkins.

Bryden, M. P. (1963) Ear preference in auditory perception. *Journal of Experimental Psychology, 16,* 291-299.

Buckwald, J. S., & Hung, C. M. (1975). Far-field acoustic response: Origins in the cat. *Science, 189,* 382-384.

Burkard, R.F., & Secor, C. (2002). Overview of auditory evoked potentials. In J. Katz (Ed.), *Handbook of clinical audiology* (5th ed., pp. 233-248). Baltimore, MA: Lippincott Williams & Wilkins

Burke, B. D. (1975). Variables affecting stutterers' initial reactions to delayed auditory feedback. *Journal of Communication Disorders, 8,* 141-155.

Burnett, T. A., Freeland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice FO responses the manipulations in pitch feedback. *The Journal of the Acoustical Society of America, 103*, 3153-3161.

Casseday, J. H., Covey, E., & Grothe, B. (1997). Neural selectivity and tuning for sinusoidal frequency modulation in the inferior colliculus fo the big brown bat, *Eptesicus fuscus. Journal of Neurophysiology, 77,* 1595-1602.

Çelik, M., Seleker, F. K., Sucu, H., & Forta, H. (2000). Middle latency auditory evoked potentials with Parkinsonism. *Parkinsonism and Related Disorders, 6,* 95-99.

Cëponiene, R., Shestakova, A., Balan, P., Alku, P., Yiaguchi, K., & Näätänen, R. (2001). Children's auditory event-related potentials index sound complexity and "speechness." *International Journal of Neuroscience, 109,* 245-260.

Chan, Y.W., Woo, E. K.W., Hammond, S. R., Yiannikas, C., & McLeod, J.G. (1988). The interaction between sex and click polarity in brainstem auditory potentials evoked from control subjects of Oriental and Caucasian origin. *Electroencephalography and Clinical Neurophysiology, 71,* 77-80.

Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice F_0 responses to pitch-shifted voice feedback during English speech. *Journal of the Acoustical Society of America, 121,* 1157-1163.

Chermak, G. D. (2001). Auditory processing disorder: An overview for the clinician. *The Hearing Journal, 54(7),* 10-25.

Chermak, G. D., & Musiek, F. E. (1997) *Central auditory processing disorders: New perspectives.* San Diego, CA: Singular Publishing Group, Inc.

Cherry, E., & Sayers, B. (1956). Experiments upon total inhibition of stammering by external control and some clinical results. *Journal of Psychomotor Research, 1,* 233-246.

Chiappa, K. H., Gladstone, K. J., & Young, R. R. (1979). Brainstem auditory evoked responses: Studies of waveform variations in 50 normal human subjects. Archives *of Neurology, 36,* 81-87.

Chung, G. H., Han, Y. M., Jeong, S. H., & Jack, Jr., C. R. (2005). Functional heterogeneity of the supplementary motor area. *American Journal of Neuroradiology, 26*, 1819-1823.

Civier, O., & Guenther, F. H. (2005). Simulations of feedback and feedforward control in stuttering. Proceedings of the 7th Oxford Dysfluency Conference, St. Catherine's College, Oxford University, 29th June - 2nd July 2005.

Coats, A. C. (1978). Human auditory nerve action potentials in brainstem evoked responses: Latency intensity functions in detection of cochlear and retrocochlear abnormality. *Archives of Otolaryngology, 104,* 709-717.

Coats, A. C., & Martin, J. L. (1977). Human auditory nerve action potentials and brainstem evoked responses: Effect of audiogram shape and lesion location. *Archives of Otolaryngology, 103,* 605-622.

Cooper, B. G., & Goller, F. (2004). Partial muting leads to age-dependent modification of motor patterns underlying crystallized zebra finch song. *Journal of Neurobiology, 61*, 317-332.

Coren, S., & Porac, C. (1977). Fifty centuries of right-handedness: The historical record. *Science, 198,* 631-632.

Corey, D. M., & Cuddapah, V. A. (2008). Delayed feedback effects during reading and conversation tasks: Gender differences in fluent adults. *Journal of Fluency Disorders, 33,* 291-305.

Cornelisse, L. E., Gagné, J. P., & Seewald, R. C. (1991). Ear level recordings of the long-term average spectrum of speech. *Ear and Hearing, 12,* 47-54.

Courchesne, E. (1978). Changes in P3 waves with event repetition: long-term effects on scalp distribution and amplitude. *Electroencephalography and Clinical Neurophysiology, 45,* 754-766.

Crapse, T. B., & Sommer, M.A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience, 9,* 587-600.

Crottaz-Herbette, S., & Ragot, R. (2000). Perception of complex sounds: N1 latency codes pitch and topography codes spectra. *Clinical Neurophysiology, 111,* 1759-1766.

Crowley, K.E., & Colrain, I.M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep, and modality. *Clinical Neurophysiology, 115,* 734-744.

Curio, G., Neuloh, G., Numminen, J., Jousmaki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping,* 9, 183-191.

Curry, F. K. W. (1967). A comparison of left-handed and right-handed subjects on verbal and non-verbal dichotic listening tasks. *Cortex, 3,* 343-352.

Davis, H. (1964). Enhancement of evoked cortical potentials in humans related to a task requiring a decision. *Science, 145,* 182-183.

Davis, H. (1965). Slow cortical responses evoked by acoustic stimuli. *Acta Oto-laryngologica (Supplement), 59,* 179-185.

Davis, H., Mast, T., Yoshie, N., & Zerlin, S. (1966). The slow response of the human cortex to auditory stimuli: Recovery process. *Electroencephalography and Clinical Neurophysiology, 21*, 105-133.

Davis, H., & Zerlin, S. (1966). Acoustic relations of the human vertex potential. *Journal of the Acoustical Society of America, 39*, 109-116.

Dayalu, V. N., Kalinowski, J., & Stuart, A. (2005). Stuttering frequency on meaningful and nonmeaningful words in adults who stutter. *Folia Phoniatrica, 57,* 193-201.

Dean, M. S., & Martin, F. M. (2000). Insert earphone depth and the occlusion effect. *American Journal of Audiology, 9,* 1-4.

Dehan, C., & Jerger, J. (1990). Analysis of gender differences in the auditory brainstem response. *Laryngoscope, 100,* 18-24.

Deiber, M.P., Ibanez, V., Fischer, C., Perrin, F., & Mauguiere, F. (1988). Sequential mapping favours the hypothesis of distinct generators for Na and Pa middle latency auditory evoked potentials. *Electroencephalography and Clinical Neurophysiology, 71,* 187-197.

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

Desmedt**,** J. E., & Debecker, J. (1979). Wave form and neural mechanism of the decision P350 elicited without pre-stimulus CNV or readiness potential in random sequences of near-threshold auditory clicks and finger stimuli. *Electroencephalography and Clinical Neurophysiology, 47,* 648-670.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti., G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91,* 176-180

Diesch, E., Eulitz, C., Hampson, S., & Ross, B. (1996). The neurotopography of vowels as mirrored by evoked magnetic field measurements. *Brain & Language, 53,* 143-168.

Digeser, F.M., Wohlberedt, T., & Hoppe, U. (2009). Contribution of spectrotemporal features on auditory event-related potentials elicited by consonant-vowel syllables. *Ear and Hearing, 30,* 704-712.

Dirks, D.D. (1964). Perception of dichotic and monaural verbal material and cerebral dominance for speech. *Acta Otolaryngology, 58,* 73-80.

Don, M., & Kwong, B. (2002). Auditory brainstem response: Differential diagnosis. In J. Hall (Ed.), *Handbook of clinical audiology* (5th ed., pp. 274-297). Baltimore, MA: Lippincott Williams & Wilkins.

Don, M., Ponton, C.W., Eggermont, J.J., & Masuda, A. (1993). Gender differences in cochlear response time: An explanation for gender amplitude differences in the unmasked auditory brain-stem response. *Journal of the Acoustical Society of America, 94*, 2135-2148.

Donchin, E. (1981). Surprise!...Surprise?. *Psychophysiology, 18,* 493-513.

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Science, 11,* 357-374.

Doupe, A.J., & Kuhl, P.K. (1999). Birdsong and human speech common themes and mechanisms. *Annual Review of Neuroscience,* 22, 567–631.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature, 384,* 159-161.

Dronkers, N.F., Redfern, B.B. & Knight, R.T. (2000). The neural architecture of language disorders. In M.S. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (pp. 949-958). Cambridge: The MIT Press.

Dronkers, N.F., Shapiro, J.K., Redfern, B. & Knight, R.T. (1992). The role of Broca's area in Broca's aphasia. *Journal of Clinical and Experimental Neuropsychology, 14,* 52-53.

Duncan, J., & Katz, J. (1983). Language and auditory processing: Top down plus bottom up. In E. Lasky & J. Katz (Eds.), *Central auditory processing disorders: Problems of speech, language, and learning* (pp. 31-45). Baltimore, MD: University Park Press.

Duncan-Johnson, C. C. (1981). P300 latency: A new metric of information processing. *Psychophysiology, 18,* 207-215

Durrant, J. D. (1990). Extratympanic electrode support via vented earmold. *Ear and Hearing, 11,* 468-469.

Durrant, J. D., Sabo, D. L., & Hyre, R. J. (1990). Gender, head size, and ABRs examined in large clinical sample. *Ear and Hearing, 11,* 210-214.

Dyson, B. J., & Alain, C. (2004). Representation of concurrent acoustic objects in primary auditory cortex. *Journal of the Acoustical Society of America, 115*, 280-288.

Eddins, A. C., & Peterson, J. R. (1999). Time-intensity trading in the late auditory evoked potential. *Journal of Speech, Language and Hearing Research, 42,* 516-525.

Efron, R., & Yund, E. W. (1975). Dichotic competition of simultaneous tone burst of different frequency. III. The effect of stimulus parameters on suppression and ear dominance functions. *Neuropsychologia, 13,* 151-161.

Egan, J.P. (1948). Articulation testing methods. *Laryngoscope, 58,* 955-991.

Eggermont, J. J., & Ponton, C. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: correlations with changes in structure and speech perception. *Acta Oto-laryngologica, 123,* 249-250.

Elberling, C., Bak, C., Kofoed, B., Lebech, J., & Saermark, K. (1980). Magnetic auditory responses for the human brain. A preliminary report. *Scandinavian Audiology, 9,* 185-190.

Elberling, C., & Parbo, J. (1987). Reference data of ABRs in retrocochlear diagnosis. *Scandinavian Audiology, 16,* 49-55.

Eldert, M. A., & Davis, H. (1951). The articulation function of patients with conductive deafness. *Laryngoscope, 61*, 891-909.

Elfner, L.F., Gustafson, D.J., & Williams, K.N. (1976). Signal onset and task variables in auditory evoked potentials. *Biological Psychology, 4,* 197-206.

Elidan, J., Sohmer, H., Gafni, M., & Kahana, E. (1982). Contribution of changes in click rate and intensity on diagnosis of multiple sclerosis. *Acta Neurologica Scandinavian, 65*, 570-585.

Elman, J.L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *Journal of Acoustic Society of America, 70,* 45-50.

Emerson, R. W., Johnson, G. M., Bosco, R. A. (2010). Social Aims. In G. M. Johnson, R. A. Bosco, & J. Myerson (Eds.). *The collected works of Ralph Waldo Emerson, Volume VIII: Letters and social aims* (p.50). Harvard University Press.

Eulitz, C., Diesch, C., Pantev, S., Hampson, & Elbert, T. (1995). Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *Journal of Neuroscience, 15,* 2748-2755.

Fabiani, M., Sohmer, H., Tait, C., Gafni, M., & Kinarti, R. (1979). A functional measure of brain activity. *Electroencephalography and Clinical Neurophysiology, 47,* 483-491.

Fairbanks, G., Everitt, W., & Jerger, R. (1954). Methods for time or frequency compression-expansion of speech. *Trans IRE-PGA, AU-2,* 7-12.

Fairbanks, G., & Guttman, N. (1958). Effects of delayed auditory feedback. *Journal of Speech and Hearing Research, 1*, 12-22.

Feinberg, I. (1978). Efference copy and corollary discharge: implications for thinking and its disorders. *Schizophrenia Bulletin, 4,* 636-640.

Ferrandez, A.M., Hugueville, L., Lehéricy, S., Poline, J.B., Marsault, C., & Pouthas, V. (2003). Basal ganglia and supplementary motor area subtend duration perception: An fMRI study. *NeuroImage, 19,* 1532-1544.

Ferrari, P.F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17,* 1703-1714.

Finley, D. J. (1952). *Statistical method in biological essay.* Cambridge University Press.

Fitzgerald, P.G., & Picton, T.W. (1983). Event-related potentials recorded during the discrimination of improbable stimuli. *Biological Psychology, 17*, 241-276.

Florentine, M., Fastl, H., & Buus, S. (1988). Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking. *Journal of the Acoustical Society of America, 84,* 195-203.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-Mental State" A practical method for grading the cognitive state of patient of the clinician. *Journal of Psychiatry Research, 12,* 189-198.

Ford, J.M., & Mathalon, D.H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *Journal of Psychiatric Research, 38*, 37-46.

Ford, J.M., Mathalon, D.H., Heinks, T., Kalba, S., & Roth, W.T. (2001). Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry, 158*, 2069-2071.

Ford, J. M., Mathalon, D. H., Kalba, S., Whitfield, S., Faustman, W.O., & Roth, W. T. (2001). Cortical responsiveness during inner speech in schizophrenia: An event-related potential study. *American Journal of Psychiatry, 158*, 1914-1916.

Ford, J. M., Mathalon, D. H., Whitfield, S., Faustman, W.O., & Roth, W. T. (2002). Reduced communication between frontal and temporal lobes during talking in schizophrenia. *Biology Psychiatry, 51,* 485-492.

Ford, J. M., Mohs, R. C., Pfefferbaum, A., Kopell, B. S. (1980). On the utility of P3 latency and RT for studying cognitive processes. *Progress in Brain Research, 54,* 661-667.

Fraser, G. R. (1964), Sex-linked recessive congenital deafness and the excess in males in profound childhood deafness. *Annals of Human Genetics, 29,* 171-196.

Friston, K. J., & Frith, C. D. (1995). Schizophrenia: A disconnection syndrome? *Clincial Neuroscience, 3,* 89-97.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.

Frith, C., & Done, J. (1989). Positive symptoms of schizophrenia. *British Journal of Psychiatry, 154,* 569-570.

Fruhstorfer, H., Soveri, P., & Jävilehto, T. (1970). Short-term habituation of the auditory evoked response in man. *Electroencephalography and Clinical Neurophysiology, 28,* 153-161.

Fukawa, T., Yoshioka, H., Ozawa, E., & Yoshida, S. (1988). Difference of susceptibility to delayed auditory feedback between stutterers and nonstutterers. *Journal of Speech and Hearing Research, 31,* 475-479.

Fujita, A., Hyde, M. L., and Alberti, P. W. (1991). ABR latency in infants: Properties and applications of various measures. *Acta Oto-laryngologica, 111,* 53-60.

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

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119,* 593-609.

Garner, W. R., & Miller, G. A. (1947). The masked threshold of pure tones as a function of duration. *Journal of Experimental Psychology, 37,* 293-303.

Gazzaniga, M., & Sperry, R. (1967). Language after section of the cerebral commissures. *Brain, 90,* 131-148.

Geisler, M. W., & Murphy, C. (2000). Event-related brain potentials to attended and ignored olfactory and trigeminal stimuli. *International Journal of Psychophysiology, 37,* 309-315.

Gelfand, S. A. (2002). The acoustic reflex. In J. Katz (Ed.), *Handbook of clinical audiology* (pp.205-232). New York: Lippincott Williams & Wilkins.

Gerin, P., Pernier, J., & Peronnet, F. (1972). Amplitude of acoustic averaged evoked potentials of the vertex intensity of the stimuli. *Brain Research, 36,* 89-100.

Gloor, P. (1997). *The temporal lobe and limbic system.* NY: Oxford University Press.

Godey, B., Schwartz, D., de Graaf, J. B., Chauvel, P., & Liegeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clinical Neurophysiology, 112,* 1850-1859.

Goff, W. R. (1978). The scalp distribution of auditory evoked potentials. In R. F. Naunton, & C. Fernandez (Eds.), *Evoked electrical activity in the auditory nervous system* (pp. 505-524). New York: Academic Press, Inc.

Goff, W. R., Alison, T., & Vaughan, Jr., H. G. (1978). The functional neuro-anatomy of event-related potentials. In E. Callaway, E. Tueting, & S. H. Koslow (Eds.), *Event-related potentials in man* (pp. 1-79). New York: Academic Press.

Golgeli, A., Suer, C., Ozesmi, N., Dolu, M., Ascioglu, M., & Sahin, O. (1999). The effect of sex differences on event-related potentials in young adults. *International Journal of Neuroscience, 99,* 69-77.

Goodin, D. S., Squires, K. C., & Starr, A. (1983). Variations in early and late event-related components of the auditory evoked potential with task difficulty. *Electroencephalography and Clinical Neurophysiology, 55,* 680-686.

Goodman, A. C. (1965). Reference zero levels for pure-tone audiometers. *ASHA, 7,* 262-263.

Green, D. M., Birshall, T. G., & Tanner, W. P. (1957). Signal detection as a function of signal intensity and duration. *Journal of the Acoustical Society of America, 29,* 523-531.

Greenlee, J. D. W., Oya, H., Kawasaki, H., Volkov, I. O., Kaufman, O. P., Kovach, C., Howard, M. A., & Brugge, J. F. (2004). A functional connection between inferior frontal gyrus and orofacial motor cortex in human. *Journal of Neurophysiology, 92,* 1153-1164.

Griffiths, S. K., Chambers, R. D., & Bilger, R. C. (1989). Statistical dependence among the wave latencies of the auditory brain stem response. *Ear and Hearing, 10,* 299-303.

Grinnell, A. D. (1963). Neurophysiology of audition in bats: Intensity and frequency parameters. *Journal of Physiology, 167,* 38-66.

Guenther, F. H. (1994). A neural network of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72,* 43-53.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102,* 594-621.

Guenther, F. H. (2001). Neural modeling of speech production. In Proceedings of the 4th international Nijmegen speech motor conference.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*, 350-365.

Guenther, F. H. (2007). Neuroimaging of normal speech production. In R. Ingham (Ed.), *Neuroscience research in communication sciences and disorders* (pp. 1-51). San Diego: Plural.

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

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

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105,* 611-633.

Guenther, F. H., & Vladusich, T. (2009). A neural theory of speech acquisition and production. *Journal of Neurolinguistics* (in press).

Guilhoto, L. M., Quintal, V. S., & da Costa, M. T. (2003). Brainstem auditory evoked response in normal term neonates. *Arquivos de Neuro-Psiquiatria, 61*, 906-908.

Gunji, A., Hoshiyama, M., Kakigi, R. (2000). Identification of auditory evoked potentials of one's own voice. *Clinical Neurophysiology, 111,* 214-219.

Guthrie, B., Porter, J., & Sparks, D. (1983). Corollary discharge provides accurate eye position information to the oculomotor system. *Science, 221*, 1193-1195.

Hall, J. W. (1992). *Handbook of auditory evoked responses.* Needham Heights, MA: Allyn and Bacon.

Hall, J. W. (2007). *New Handbook of auditory evoked responses.* Boston, MA: Pearson Education, Inc.

Hall, J. W., Buss, E., & Grose, J. H. (2007). The binaural temporal window in adults and children. *The Journal of the Acoustical Society of America, 121,* 401-410.

Hallpike, C. S. (1965). Clinical otoneurology and its contributions in theory and practice. *Proceedings of the Royal Society of Medicine, 58,* 185-196.

Halgren, E., Squires, N. K., Wilson, C. L., Rohrbaugh, J. W., Babb, T. L., & Crandall, P.H. (1980). Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science, 210,* 803-805.

Halpin, C. (2002). The tuning curve in clinical audiology. *American Journal of Audiology, 11,* 56-64.

Halsband, U., Ito, N., Tanji, J., & Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain, 116,* 243-266.

Hargrave, S., Kalinowski, J., Stuart, A., Armson, J., & Jones, K. (1994). Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *Journal of Speech and Hearing Research, 37*, 1313-1319.

Hari, R., Aittoniemi, K., Jarvinen, M. L., Katila, T., & Varpula, T. (1980). Auditory evoked transient and sustained magnetic fields of the human brain localization of neural generators. *Experimental Brain Research, 40*, 237-240.

Hari, R., Hämäläinen, H., Hämäläinen, M., Kekoni, J., Sams, M., & Tiihonen, J. (1990). Separate finger representations at the human second somatosensory cortex. *Neuroscience, 37,* 245-249.

Hari, R., Pelizzone, M. Makela, J. P., Hallstrom, J., Leinonen, L., & Lounasmaa, O. V. (1987). Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology, 26,* 31-43.

Harrison, J., Buchwald, J., & Kaga, K. (1986). Cat P300 present after primary auditory cortex ablation. *Electroencephalography and Clinical Neurophysiology, 63,* 180-187.

Hartridge, H. (1945). Acoustical control in the flight of bats. *Nature, 156,* 490-494.

Hartsuiker, R. J., Pickering, M.J., & de Jong, N. H. (2005). Semantic and phonological context effects in speech error repairs. *Journal of Experimental Psychology, 31*, 921-932.

Hashimoto, I., Ishiyama, Y., Yoshimoto, R., & Nemoto, S. (1981). Brainstem auditory evoked potentials recorded directly from human brainstem and thalamus. *Brain, 104,* 841-859.

Hashimoto, Y., Sakai, K.L., 2003. Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. *Human Brain Mapping, 20,* 22-28.

Hawks, J. W., & Miller, J. D. (1995). A formant bandwidth estimation procedure for vowel synthesis. *The Journal of the Acoustical Society of America, 97*, 1343-1345.

He, S. Q., Dum, R. P., & Strick, P. L. (1995). Topographic organization of corticospinal projections from the frontal lobe: Motor areas on the medial surface of the hemisphere. *Journal of Neuroscience, 15,* 3284-3306.

Heiligenberg, W. (1969). The effect of stimulus chirps on a cricket's chirping (*Acheta domesticus)*. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 65,* 70-97.

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology, 42*, 180-190.

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Archives of General Psychiatry, 64,* 286-296.

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport, 17*, 1375-1379.

Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and sentence comprehension in the left temporal cortex. *Brain, 121,* 1133-1142.

Henson, Jr., O. W. (1965). The activity of the middle-ear muscles in echo-locating bats. *Journal of Physiology, 180,* 871-887.

Hikosaka, O., Sakai, K., Miyauchi, S., Takino, R., Sasaki, Y., & Pütz, B. (1996). Activation of human presupplementary motor area in learning of sequential procedures: A functional MRI study. *Journal of Neurophysiology, 76,* 617-621.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science, 182,* 177-180.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., & Konishi, J. (1996). Cortical processing mechanism while vocalizing visually presented languages. *Neuroreport, 8*, 363-367.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., & Konishi, J. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *Neuroreport, 8*, 2379-2382.

Hirsh, I. J. (1948). The influence of interaural phase on interaural summation and inhibition. *Journal of the Acoustical Society of America, 20,* 536-544.

Hirsh, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America, 31,* 759-767.

Hood, J. D. (1969). Basic audiological requirements in neur-otology. *Journal of Laryngology and Otology, 83,* 695-711.

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *The Journal of Cognitive Neuroscience, 14*, 1125-1138.

Howell, P. (1990). Changes in voice level caused by several forms of altered feedback in fluent speakers and stutters. *Language and Speech, 33,* 325-338.

Howell, P., & Powell, D. J. (1984). Hearing your own voice through bone and air: Implications for explanations of stuttering behavior from studies of normal speakers. *Journal of Fluency Disorders, 9,* 247-264.

Howell, P., & Sackin, S. (2002). Timing interference to speech in altered listening conditions. *The Journal of the Acoustical Society of America, 111*, 2842-2852.

Howell, P., El-Yaniv, N., & Powell, D. J. (1987). Factors affecting fluency in stutterers. In H. F. M. Peters, & W. Hulstijin (Eds.), *Speech motor dynamics in stuttering* (pp. 361-369). New York: Springer-Verlag.

Hoy, R. R., & Robert, D. (1996). Tympanal hearing in insects. *Annual Review of Entomology, 41*, 433-450.

Hudgins, C. V., Hawkins, J. E., Karlin, J. E., & Stevens, S. S. (1947). The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope, 57,* 52-89.

Hyde, M. (1997). The N1 response and its applications. *Audiology and Neurotology, 2,* 281-307.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reveiws Neuroscience, 7,* 942-951.

Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seitz, R.J., et al. (2001). A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 5933-5936.

Iyengar, S., & Bottjer, S. W. (2002). The role of auditory experience in the formation of neural circuits underlying vocal learning in zebra finches. *Journal of Neuroscience, 22*, 946-958.

Jacobsen, J. T., Novotny, G. M., & Elliot, S. (1980). Clinical considerations in the interpretation of auditory brainstem response audiometry. *Journal of Otolaryngology, 9,* 493-504.

Jacobson, G. P. (1994). Magnetoencephalographic studies of auditory system function. *Journal of Clinical Neurophysiology, 11*, 343-364.

Jacobson, G. P., Ahmad, B. K., Moran, J., Newman, C. W., Wharton, J., Tepley, N. (1992). N1M and P2M components in a sample of normal subjects. *Ear and Hearing, 13,* 387-395.

Jacobson, G. P., & Newman, C. W. (1990). The decomposition of the middle latency auditory evoked potential (MLAEP) Pa component into superficial and deep source contributions. *Brain Topography, 2,* 222-236.

Jasper, H. H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology, 10*, 370-371.

Jeannerod, M. (1998). *The neural and behavioral organization of goal-directed movements.* Oxford: Oxford University Press.

Jerger, J. (1973). Diagnostic audiometry. In J. Jerger (Ed.), *Modern developments in audiology* (pp.75-112). New York: Academic Press.

Jerger, J. (1997). Functional asymmetries in the auditory system. *Annals of Otology, Rhinology, and Laryngology, 106,* 23-30.

Jerger, J., Burney, P., Maudlin, L., & Crump, B. (1974). Predicting hearing loss from the acoustic reflex. *Journal of Speech and Hearing Disorders, 39*, 11-22.

Jerger , J., & Hall, J. W., III. (1980). Effects of age and sex on auditory brainstem response (ABR). *Archives of Otolaryngology, 106,* 387-391.

Jerger, J., & Jerger, S. (1971). Diagnostic significance of PB words functions. *Archives of Otolaryngology, 93,* 573-580.

Jerger, J., & Jerger, S. W. (1974). Auditory findings in brainstem disorders. *Archives of Otolaryngology, 99,* 342-349.

Jerger, J., & Jerger, S.W. (1975). Clinical validity of central auditory tests. *Scandinavian Audiology, 4,* 147-163.

Jerger, J., & Mauldin, L. (1978). Prediction of sensorineural hearing level from the brainstem evoked response. *Archives of Otolaryngology, 104,* 456-461.

Jerger, J., & Musiek, F. (2000). Report of the consensus on the diagnosis of auditory processing disorders in school-aged children. *Journal of the American Academy of Audiology, 11,* 467-474.

Jerger, J., Speaks, C., & Trammel, J. (1968). A new approach to speech audiometry. *Journal of Speech and Hearing Disorders, 33,* 318-328.

Jones, M. D. R., & Dambach M**. (**1973**).** Response to sound in crickets without tympanal organs (*Gryllus campestris L.*). *Journal of Comparative Physiology A, 87,* 89-98.

Kalinowski, J., Armson, J., Roland-Mieszkowski, M., Stuart, A., & Gracco, V. (1993). The effects of alterations in auditory feedback on stuttering frequency. *Language and Speech, 36*, 1-16.

Kalinowski, J., Stuart, A., Sark, S., & Armson, J. (1996). Stuttering amelioration at various auditory feedback delays and speech rates. *European Journal of Disorders of Communication, 31*, 259-269.

Katz, J. (1962). The use of staggered spondaic words for assessing the integrity of the central auditory nervous system. *Journal of Auditory Research, 2,* 327-337.

Katz, J., Smith, P., & Kurpita, B. (1992). Categorizing test findings in children referred for auditory processing deficits. *SSW Reports, 14,* 1-6.

Katz, J., & Ivey, R.G. (1994). Spondaic procedures in central testing. In J. Katz (Ed.), *Handbook of clinical audiology* (4th ed., pp. 239-255). Baltimore, MD: Williams & Wilkins.

Katahira, K., Abla, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: An ERP study. *Neuroscience Research, 61,* 120-128.

Kawato, M. & Wolpert, D. (1998). Internal models for motor control. *Norvartis Foundation Symposium 218. Sensory guidance of movement* (pp. 291-307).

Keating, L. W., & Ruhm, H. B. (1971). Some observations on the effects of attention to stimuli on the amplitude of the acoustically evoked response. *International Journal of Audiology, 10,* 177-184.

Keidel, W. D., & Spreng, M. (1965). Neurophysiological evidence for the Stevens power function in man. *Journal of the Acoustical Society of America, 38,* 191-195.

Keith, R. W. (1977). Synthetic sentence identification test. In R.W. Keith (Ed.), *Central auditory dysfunction* (pp.73-102). New York: Grune & Stratton.

Keith, R. W. (1981). Audiological and auditory language tests of central auditory function. In R. W. Keith (Ed.), *Central auditory and language disorders in children* (pp. 61-76). Houston, TX: College-Hill Press.

Kent, R. D. (2000). Research on speech motor control and its disorders: A review and prospective. *Journal of Communication Disorders, 33,* 391-428.

Keppel, G., & Wickens, T. D. (2004). *Design and Analysis: A researcher's handbook (4th ed.).* Upper Saddle River, NJ: Pearson Education.

Kertesz, A. (1994). Neuropsychological evaluation of language. *Journal of Clinical Neurophysiology, 11*, 205-215.

Kileny, P. R., & Kripal, J. P. (1987). Test-retest variability of auditory event-related potentials. *Ear and Hearing, 8,* 110-114.

Kilney, P.R., Paccioretti, D., & Wilson, A. F. (1987). Effects of cortical lesions on middle-latency auditory evoked responses (MLR). *Electroencephalography and Clinical Neurophysiology, 66,* 108-120.

Kimura, D. (1961a). Cerebral dominance and the perception of verbal stimuli. *Candiaian Journal of Psychology, 15,* 166-171.

Kimura, D. (1961b). Some effects of temporal-lobe damage on auditory perception. *Candiaian Journal of Psychology, 15,* 156-165.

Knight, R.T., Hillyard, S.A., Woods, D.L. & Neville, H.J. (1980). The effects of frontal and tempo-parietal lesions on the auditory evoked potentials in man. *Electroencephalography and Clinical Neurophysiology, 50,* 112-124.

Koch, U., & Grothe, B. (2000). GABAergic and glycinergic inhibition sharpens tuning for frequency modulations in the inferior colliculus of the big brown bat. *Journal of Neurophysiology, 80,* 71-82.

Kohler, E., Keysers, C., Umilta, M. A. Fogassi, L., Gallese, V., Rizzolatti. G. (2002). Hearing sounds, understanding actions: action representation in Mirror Neurons. *Science, 297,* 846-848

Konishi, M. (1965a). Effects of deafening on song development in American Robins and Black-headed Grosbeaks *Zeitschrift für. Tierpsychologie, 22,* 584-599.

Konishi, M. (1965b). The role of auditory feedback in the control of vocalization in the White-crowned Sparrow. *Zeitschrift für. Tierpsychologie, 22*, 770-783.

Konishi, M. (2004). The role of auditory feedback in Birdsong. *Annals of the New York Academy of Sciences, 1016,* 463-475.

Konkle, D.F., & Rintelmann, W.F. (1983). Masking in speech audiometry. In D.F. Konkle & W.F. Rintelmann (Eds.), *Principles of speech audiometry* (pp. 285-319). Baltimore, MD: University Park Press.

Knight, R. T., Scabini, D., Woods, D. L., & Clayworth, C. C. (1989). Contributions of temporal-parietal junction to the human auditory P3. *Brain Research, 502,*109-116.

Kraus, N., & McGee, T. (1993). Clinical implications of primary and nonprimary pathway contributions to the middle latency response generating system. *Ear and Hearing, 14*, 36-48.

Kraus, N., & McGee, T. (1995). The middle latency response generating system. *Electroencephalography and Clinical Neurophysiology, 44 (Suppl.), 93-101.*

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

Kraus, N., McGee, T., Carrell, T., Sharma, A., Micco, A., & Nicol, T. (1993). Speech-evoked cortical potentials in children. *Journal of the American Academy of Audiology, 4,* 238-248.

Kraus, N., Smith, D., Reed, N. L., Stein, L. K., & Cartee, C. (1985). Auditory middle latency responses in children: Effects of age and diagnostic category. *Electroencephalography and Clinical Neurophysiology, 62,* 343-351.

Kroodsma, D. E., & Konishi, M. (1991). A suboscine bird (Eastern Phoebe, Sayornis phoebe) develops normal song without auditory feedback. *Animal Behavior, 42*, 477-488.

Kudo, N., Nakagome, K., Kasai, K., Araki, T., Fukuda, M., Kato, N., & Iwanami, A. (2004). Effects of corollary discharge on event-related potentials during selective attention task in healthy men and women. *Neuroscience Research, 48,* 59-64.

Kuhl, P. K. (2000). A new view of language acquisition*. Proceedings of the National Academy of Sciences, 97,* 11850-11857.

Kurdziel, S. A., Noffsinger, P.D., & Olsen, W. (1976). Performance by cortical lesion patients on 40 and 60 percent time-compressed materials. *Journal of the American Audiological Society, 2,* 3-7.

Kuriki, S., Mori, T., & Hirata, Y. (1999). Motor planning center for speech articulation in the normal human brain. *NeuroReport, 10,* 765-769.

Kurtzberg, D. (1989). Cortical event-related potential assessment of auditory system function. *Seminars in Hearing, 10*, 252-261.

Kurtzberg, D., Vaughan, H. G. Jr., Kreuzer, J. A. (1979). Task related cortical potentials in children. *Progress in Clinical Neurophysiology, 6,* 216-223.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science, 197,*792-795.

Lamandella, J. T. (1977). The limbic system in human communication. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolingusitics* (Vol 3, pp. 157-222). NY: Academic Press.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research, 14*, 677-709.

Larson, C. R. (1988). Brain mechanisms involved in the control of vocalization. *Journal of Voice, 2,* 301-311.

Larson, C. R., Burnett, T. A., Kiran, S., & Hain, T. C. (2000). Effects of pitch-shift velocity on voice Fo responses. *Journal of the Acoustical Society of America, 107*, 559-564.

Lazorthes, G., LaComme, Y., Ganbert, J., & Planel, H. (1961). Composition of the auditory nerve. *Presse Medicale, 69,* 1067-1068.

Leahy, R. M., Mosher, J. C., Spencer, M. E., Huang, M. X., & Lewine, J. D. (1998). A study of dipole localization accuracy for MEG and EEG using a human skull phantom. *Electroencephalography and Clinical Neurophysiology, 107,* 159-173.

Lee, B. S. (1950). Effects of delayed speech feedback. *Journal of the Acoustical Society of America, 22,* 824-826.

Lee, Y. S., Lueders, H., Dinner, D. S., Lesser, R. P., Hahn, J., & Klem, G. (1984). Recording of auditory evoked potentials in man using chronic subdural electrodes. *Brain, 107,* 115-131.

Leonardo, A., & Konishi, M. (1999). Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature, 399,* 466-470.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41-104.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-75.

Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29, 117-123.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 4*, 431-461.

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

Liberman, A. M. & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 243,* 489-494.

Licklider, J.C.R. (1948). The influence of interaural phase relations upon the masking of speech by white noise. *Journal of the Acoustical Society of America, 20,* 150-159.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology, 92,* 204-214.

Luppino, G., & Rizzolatti, G. (2000). The organization of the frontal motor cortex. *New in Physiological Sciences, 15,* 219-224.

Lynn, G. W., & Gilroy, J. (1972). Neuro-audiological abnormalities in patients with temporal lobe tumors. *Journal of Neurological Science, 17,* 167-184.

Lynn, G. W., & Gilroy, J. (1975). Effects of brain lesions on the perception of monotic and dichotic speech stimuli. In H. Sullivan (Ed.), *Proceedings of a Symposium on Central Auditory Processing Disorders* (pp. 47-83). Omaha, NE: University of Nebraska Medical Center.

Lynn, G. E., Gilroy, J., Taylor, P.C., & Leiser, R. P. (1981). Binaural masking level differences in neurological disorders. *Archives of Otolaryngology, 107,* 357-362.

Maeda, S. (1990). Compensatory articulation during speech: evidence from the analysis and synthesis of vocal tract shapes using an articulatory model. In: W.J. Hardcastle and A. Marshal (Eds.), *Speech production and speech modeling* (pp. 131-149). Boston, MA: Kluwer Academic Publishers.

Magliero, A., Bashore, T. R., Coles, M. G. H., & Donchin, E. (1984). On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology, 21,* 171- 186. .

Mäkelä, A.M., Alku, P., & Tiitinen, H. (2003). The auditory N1m reveals the left-hemispheric representation of vowel identity in humans. *Neuroscience Letters, 353,* 111-114.

Maraist, J. A., & Hutton, C. (1957). Effects of auditory masking upon the speech of stutterers. *Journal of Speech and Hearing Disorders, 22,* 385-389.

Margolis, R.H., & Hunter, L.L. (2000). Acoustic immittance measurements. In R.J. Roeser, M. Valente, & H. Hosford-Dunn, (Eds.), *Audiology diagnosis* (pp. 381-423). New York: Thieme.

Martin, B., A., & Boothroyd, A. (1999). Cortical, auditory, event-related potentials in response to periodic and aperiodic stimuli with the same spectral envelope. *Ear and Hearing, 20,* 33-44

Martin, B. A., Sigal, A., Kurtzberg, D., & Stapells, D. R. (1997). The effects of decreased audibility produced by high-pass noise masking on cortical event-related potentials to speech sounds/ba/ and /da/. *Journal of the Acoustical Society of America, 101,* 1585-1599.

Martin, B. A., & Stapells, D. R. (2005). Effects of low-pass noise masking on auditory event-related potentials to speech. *Ear and Hearing, 26,* 195-213.

Martin, B. A., Tremblay, K. L., & Korczak, P. (2008). Speech evoked potentials from the laboratory to the clinic. *Ear and Hearing, 29,* 285- 313.

Mattingly, I. G. & Liberman, A. M. (1988). Specialized perceiving systems for speech and other biologically significant sounds. In G. M. Edelman, W. E. Gall, and W. M. Cowan (Eds.), *Functions of the Auditory System*. (pp. 775-793). New York: Wiley.

Matzker, J. (1959). Two new methods for the assessment of central auditory function in cases of brain disease. *Annals of Otology, Rhinology, and Laryngology, 68,* 1185-1196.

Mayer, M., Dogil, G., Ackermann, H., Erb, M., Riecker, A., Wildgruber, D., et al. (2000). Prosody in speech production: A paradigm for functional imaging and first result. In *Proceedings of the fifth seminar on speech production: Models and data* (pp. 281-184). Munich: Universität Munches.

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. *Contemporary Issues in Communication Science and Disorders, 31,* 105-122.

McClelland, R. J., McCrea, R. S. (1979). Intersubject variability of the auditory-evoked brain-stem potentials. *Audiology, 18*, 462-471.

McFarland, W. H., Vivion, M. C., & Gloodstein, R. (1977). Middle components of the AER to tone-pips in normal-hearing and hearing impaired subjects. *Journal of Speech and Hearing Research, 20,* 781-798.

Medwetsky, L. (2002). Central auditory processing testing: A battery approach. In J. Katz (Ed.), *Handbook of clinical audiology* (pp. 510-524). New York: Lippincott Williams & Wilkins.

Mendel, L., & Danhauer, J. (1997). *Audiologic evaluation and management and speech perception assessment.* San Diego, CA: Singular Publishing Group, Inc.

Metzner, W. (1989). A possible neuronal basis for Doppler-shift compensation in echo-locating horseshoe bats. *Nature, 341,* 529-532.

Metzner, W. (1993). A audio-vocal interface in echolocating horse bats. *Journal of Neuroscience, 13,* 1899-1915.

Michalewski, H. J., Thompson, L. W., Patterson, J. V., Bowman, T. E., & Litzelman, D. (1980). Sex differences in the amplitudes and latencies of the human auditory brain stem potential. *Electroencephalography and Clinical Neurophysiology, 48,* 351-356.

Milner, B., Taylor, S., & Sperry, R. W. (1968). Lateralized suppression of dichotically presented digits after commissural section in man. *Science, 161,* 184-185.

Milner, B. A. (1969). Evaluaiton of auditory function by computer techniques. *International Journal of Audiology, 8,* 361-370.

Møller, A. R. (1985). Origin of latency shift of cochlear nerve potentials with sound intensity. *Hearing Research, 17,* 177-189.

Møller, A. R., & Jannetta, P. J. (1981). Compound action potentials recorded intracranially from the auditory nerve in man. *Journal of Experimental Neurology, 74,* 862-874.

Møller, A. R., & Jannetta, P. J. (1982). Evoked potentials from the inferior colliculus in man. *Electroencephalography and Clinical Neurophysiology, 53,* 612-620.

Møller, A. R., & Jannetta, P. J. (1983). Auditory evoked potentials recorded from the cochlear nucleus and its vicinity in man. *Journal of Neurosurgery, 59,* 1013-1018.

Mooney, R., & Prather, J. F. (2005). The HVC microcircuit: The synaptic basis for interactions between song motor and vocal plasticity pathways. *Journal of Neuroscience, 25*, 1952-1964.

Moore, B. C. J. (1997). An introduction to the psychology of hearing (4th edition). London: Academic Press.

Morgan, A. H., McDonald, P. J., & MacDonald, H. (1971). Differences in bilateral alpha activity as a function of experimental task, with a note on lateral eye movements and hypnotizability. *Neuropsychologia, 9,* 459-469.

Morgan, M. D., Cranford, J. L., & Burk, K. (1997). P300 event-related potentials in stutterers and nonstutterers. *Journal of Speech, Language, and Hearing Research, 40*, 1334-1340.

Moss, C. F., & Sinha, S. R., (2003). Neurobiology of echolocation in bats. *Current Opinion in Neurobiology, 13,* 755-762.

Mueller, H.G., Beck, W. G., & Sedge, R. K. (1987). Comparison of the efficiency of cortical level speech tests. *Seminars in Hearing, 8,* 279-298.

Munhall, K. G. (2001). Functional imaging during speech production. *Acta Psychologica, 107,* 95-117.

Musiek, F. E. (1983). Assessment of central auditory dysfunction: The dichotic digits test revisited. *Ear and Hearing, 4,* 79-83.

Musiek, F. E., & Chermak, G. D. (1994). Three commonly asked questions about central auditory processing disorders: Management. *American Journal of Audiology, 3,* 23-27.

Musiek, F. E., Pinherio, M. L., & Wilson, D. H. (1980). Auditory pattern perception in "split-brain" patients. *Archives of Otolaryngology, 106,* 610-612.

Musiek, F. E., & Sachs, Jr., E. (1980). Reversible neuroaudiologic findings in a case of right frontal lobe abscess with recovery. *Archives of Otolaryngology, 106,* 280-283.

Müller-Preuss, P. (1978). Single unit responses of the auditory cortex in the squirrel monkey to self-produced and loudspeaker transmitted vocalizations. *Neuroscience Letters, Supplement, 1*, S.7.

Müller-Preuss, P., Newman, J. D., & Jürgens, U. (1980). Anatomical and physiological evidence for a relationship between the 'cingular' vocalization area and the auditory cortex in the squirrel monkey. *Brain Research, 202*, 307-315.

Müller-Preuss, P., & Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain Research, 215*, 61-76.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral Brain Science, 13*, 201-288.

Näätänen, R. (1992). *Attention and brain function.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Näätänen, R. (2003). Mismatch negativity: Clinical research and possible applications. *International Journal of Psychophysiology, 48*, 179-188.

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

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology, 115*, 140-144.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*, 375-425.

Naylor, R.V. (1953). A comparative study of methods of estimating the severity of stuttering. *Journal of Speech and Hearing Disorders, 18,* 30-37.

Naylor, H., Halliday, R., Callaway, E., Yano, L., & Walton, P. (1987). P3 as an index of visual information processing. *Electroencephalography and Clinical Neurophysiology 40,* (Suppl.), 235-240.

Neijenhuis, K. A. M., Stollman, M. H. P., Snik, A. F. M., & Van den Broek, P. (2001). Development of a central auditory test battery for adults. *Audiology, 40,* 69-77.

Nelson, D. A., & Lassman, F. M. (1968). Effects of intersignal interval on the human auditory evoked response. *Journal of the Acoustical Society of America, 33,* 1529-1532.

Neuweiler, G. (2003). Evolutionary aspects of bat echolocation. *Journal of Comparative Physiology A:, 189,* 245-256.

Nolte, J. (2002). *The human brain: An introduction to its functional anatomy* (5th ed.). St. Louis, MO: Mosby Inc

Nooteboom, S. C. (1980). Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In V. A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 87-95). New York: Academic Press.

Nordeen, K. W., & Nordeen, E. J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and Neural Biology, 57,* 58-66.

Nordlund, B. (1964). Directional audiometry. *Acta Oto-laryngology 57,* 1-18.

Northern, J.L., & Gabbard, S.A. (1994). The acoustic reflex. In J. Katz (Ed.), *Handbook of clinical audiology* (4th ed., pp. 300-316). Baltimore, MD: Williams & Wilkins

Nottebohm. F. (1968). Auditory experience and song development in the chaffinch, *Fringilla coelebs. International Journal of Avian Science, 110,* 549-569.

Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters, 265*, 119-122.

O'Brien, P. J., & Stuart, A. (2001). The effect of auditory stimulus during on the P300 response. *Journal of Speech-Language Pathology and Audiology, 25,* 19 - 23.

Obleser, J., Eulitz, C., Lahiri, A., & Elbert, T. (2001). Gender differences in functional hemispheric asymmetry during the processing of vowels as reflected by the human brain magnetic response. *Neuroscience Letters, 314,* 131-134.

Obleser, J., Lahiri, A., & Eulitz, C. (2004). Magnetic brain response mirrors extraction of phonological features form spoken words. *Journal of Cognitive Neuroscience, 16,* 31-39.

Obleser, J., Rockstroh, B., & Eulitz, C. (2004). Gender differences in hemispheric asymmetry of syllable processing: Left-lateralized magnetic N100 varies with syllable categorization in females. *Psychophysiology, 41,* 783-788.

Okada, Y. C., Kaufman, L., & Williamson, S. J. (1983). The hippocampal formation as a source of the slow endogenous potentials. *Electroencephalography and Clinical Neurophysiology, 55,* 417-426

Okanoya, K., & Yamaguchi, A. (1997). Adult Bengalese finches (Lonchura striate var. domestic) require real-time auditory feedback to produce normal song syntax. *Journal of Neurobiology, 33,* 343-356.

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburg Inventory. *Neuropsychologia, 9,* 669-680.

Onishi, S., & Davis, H. (1968). Effects of duration and rise time of tone bursts on evoked potentials. *Journal of the Acoustical Society of America, 44,* 582-591.

Ostroff, J.M., Martin, B. A., & Boothroyd, A. (1998). Cortical evoked response to acoustic change within a syllable. *Ear and Hearing, 19,* 290-297.

Paetau, R. (2002). Magnetoencephalography in pediatric neuroimaging. *Developmental Studies, 5,* 361-370.

Pantev, C., Eulitz, C., Hampson, S., Ross, B., & Roberts, L. (1996). The auditory evoked ''off'' response: Sources and comparison with the ''on'' and the ''sustained'' responses. *Ear and Hearing, 17,* 255–265.

Pantev, C., Hoke, M., Lütkenhöner, B., & Lehnertz, K. (1991). Neuromagnetic evidence of functional organization of the auditory cortex in humans. *Acta Oto-laryngologica (Supplement*), 49, 106-114.

Perkins, W. H. (1996). *Stuttering and science*. San Diego, CA: Singular Publishing Group.

Perrault, N. & Picton, T.W. (1984). Event-related potentials recorded from the scalp and nasopharynx. I. N1 and P2. *Electroencephalography and Clinical Neurophysiology, 59,* 177-194.

Peschke, C., Ziegler, W., Kappes, J., & Baumgaertner, A. (2009). Auditory-motor integration during fast repetition: The neuronal correlates of shadowing. *NeuroImage, 47*, 392-402.

Phillips, D. P. (1998). Sensory representations, the auditory cortex, and speech perception. *Seminars in Hearing, 19,* 319-332.

Phillips, D. P. (2002). Central auditory system and central auditory processing disorders: Some conceptual issues. *Seminars in Hearing, 23,* 251-261.

Picton, T. W., Alain, C., Woods, D. L., John, M. S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., & Trujillo, N. J. (1999). Intracerebral sources of human auditory-evoked potentials. *Audiology and Neurotology, 4,* 64-79.

Picton T. W., Alain, C., Otten, L., Ritter, W., & Achim A. (2000). Mismatch negativity: different water in the same river. *Audiology and Neurotology, 5,* 111-139.

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

Picton. T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials. II: Effects of Attention. *Electroencephalography and Clinical Neurophysiology, 36,* 191-199.

Picton, T. W., Woods, D. L., Baribeau-Braun, J., & Healy, T. M. G. (1977). Evoked potential audiometry. *Journal of Otolaryngology, 6,* 90-119.

Pinheiro, M. L. (1976). Auditory pattern perception in patients with right and left hemisphere lesions. *Ohio Journal of Speech and Hearing, 2,* 9-20.

Pinheiro, M. L., & Tobin, H. (1969). Interaural intensity difference for intracranial localization. *Journal of the Acoustical Society of America, 46,* 1482-1487.

Pinheiro, M. L., & Tobin, H. (1971). The interaural intensity difference as a diagnostic indicator. *Acta Oto-Laryngology, 71*, 326-328.

Plomp, R., & Bouman, A. (1959). Relation between hearing threshold and duration for tone pulses. *Journal of the Acoustical Society of America, 31,* 749-758.

Polich, J. (1989). Frequency, intensity, and duration as determinants of P300 from auditory stimuli. *Journal of Clinical Neurophysiology, 6*, 277-286.

Polich, J., Howard, L., & Starr, A. (1985). Effects of age on the P300 component of the event-related potential from auditory stimuli. *The Journal of Gerontology, 40,* 721-726.

Polich, J., & Starr, A. (1983). Middle, late, and long latency auditory evoked potentials. In B. E. Moore (Ed.), Bases of auditory brainstem evoked responses (pp. 345-361). New York: Grune & Stratton.

Ponton, C.W., Vasama, J.P., Tremblay, K., Khosla, D., & Don, M. (2000). Experience-related increases in interhemispheric correlations of evoked neurophysiological activity following profound deafness. *Hearing Research, 152,* 32-44.

Ponton, C.W., Vasama, J.P., Tremblay, K., Khosla, D., Kwong, B., & Don, M. (2001). Plasticity in the adult human central auditory system: evidence from late-onset profound unilateral deafness. *Hearing Research, 154,* 32-44.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition 77,* 97-131.

Poulet, J.F. (2005). Corollary discharge inhibition and audition in the stridulating cricket. *Journal of Comparative Physiology, 191*, 979-986.

Poulet, J.F., & Hedwig, B. (2002). A corollary discharge maintains auditory sensitivity during sound production. *Nature, 418,* 872-876.

Poulet, J.F., & Hedwig, B. (2003a). A corollary discharge mechanism modulates central auditory processing in singing crickets. *Journal of Neurophysiology, 89*, 1528-1540.

Poulet, J.F., & Hedwig, B. (2003b). Corollary discharge inhibition of ascending auditory neurons in the stridulating cricket. *Journal of Neuroscience, 23,* 4717-4725.

Poulet, J.F., & Hedwig, B. (2006). The cellular basis of a corollary discharge. *Science, 311*, 518-522.

Pritchard, W.S. (1981). Psychophysiology of P300. *Psychological Bulletin, 89,* 506-540.

Purcell. D.W. & Munhall, K.G. (2006). Adaptive control of vowel formants frequency: Evidence from real-time formant manipulation. *Journal of the Acoustical Society of America, 120,* 996-977.

Rapin, I., Schimmel, H., Tourk, L.M., Krasnegor, N.A., & Pollark, C. (1966). Evoked responses to clicks and tones of varying intensity in waking adults. *Electroencephalography and Clinical Neurophysiology, 21,* 335-344.

Rees, N.S. (1973). Auditory processing factors in language disorders: The view from Procrustes' bed. *Journal of Speech and Hearing Disorders, 38,* 304-315.

Rees, N.S. (1981). Saying more than we know: Is auditory processing disorder a meaningful concept? In R. W. Keith (Ed.), *Central auditory and language disorders in children* (pp. 94-102). Houston, TX: College-Hill Press.

Richer, F., Johnson, R.A., & Beatty, J. (1983). Sources of the late components of the brain magnetic response. *Social Neuroscience Abstracts, 9.* 656.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport, 11,* 1997-2000.

Rif, J., Hari. R., Hamalainen. M. S, & Sams. M. (1991). Auditory attention affects two different areas in the human supratemporal cortex. *Electroencephalography and Clinical Neurophysiology, 79,* 464-172.

Rizzolatti, G., & Arbib, M.A. (1998). Language within our grasp. *Trends in Neurosciences, 21,* 188-194.

Rockstroh, B., Kissler, J., Mohr, B., Eultiz, C., Lommen, U., Wienbruch, C., et al. (2001). Altered hemispheric asymmetry of auditory magnetic fields to tones and syllables in schizophrenia. *Biological Psychiatry, 49,* 694-703.

Roeser, R.J., Buckley, K.A., & Stickney, G.S. (2000). Pure tone tests. In R. J. Roeser, M. Valente, & H. Hosford-Dunn (Eds.), *Audiology diagnosis.* New York: Thieme.

Roeser, R.J., Johns, D.F., & Price, L.L. (1976). Dichotic listening in adults with sensorineural hearing loss. *Journal of the American Audiology Society, 2,* 19-25.

Roeser, R.J., Valente, M., & Hosford-Dunn, H. (2000). Diagnostic procedures in the profession of audiology. In R. J. Roeser, M. Valente, & H. Hosford-Dunn (Eds.), *Audiology diagnosis.* New York: Thieme.

Rogers, R.L., Baumann, S.B., Papanicolaou, A.C., Bourbon, T.W., Alagarsamy, S., & Eisenberg, H.M. (1991). Localization of the P3 sources using magnetoencephalography and magnetic resonance imaging. *Electroencephalography and Clinical Neurophysiology, 79,* 308-321.

Rosenhamer, H.J., Lindström, B., & Lundborg, T. (1980). On the use of click-evoked electric brainstem responses in audiological diagnosis. II. The influence of sex and age upon the normal response. *Scandinavian Audiology*, 9, 93-100.

Roth, W.T., Ford, J.M., Lewis, S.J., & Kopell, B.S. (1976). Effects of stimulus probability and task-relevance on event-related potentials. *Psychophysiology, 13,* 311-317.

Rothman, H. H., Davis, H., & Hay, I. S. (1970). Slow evoked cortical potentials and temporal features of stimulation. *Electroencephalography and Clinical Neurophysiology, 29,* 225-232.

Roup, C.M., Wiley, T.L., Safady, S.H., & Stoppenbach, D.T. (1998). Tympanometric screening norms for adults. *American Journal of Audiology, 7,* 55-60.

Rowe, M.J., III (1978). Normal variability of the brain-stem auditory evoked response in young and old adult subjects. *Electroencephalography and Clinical Neurophysiology, 44,* 459-470.

Runyan, C.M., & Adams, M. R. (1979). Unsophisticated judges' perceptual evaluations of the speech of 'successfully treated' stutterers. *Journal of Fluency Disorders, 4,* 29-48.

Runyan, C.M., Bell, J.N., & Prosek, R.A. (1990). Speech naturalness ratings of treated stutterers. *Journal of Speech and Hearing Disorders, 55,* 434-438.

Sabo, D.L., Durrant, J.D., Curtin, H., Boston, J.R., & Rood, S. (1992). Correlations of neuroanatomical measures to auditory brain stem response latencies. *Ear and Hearing, 13,* 213-222.

Salmelin, R., Hari, R., Lounasmaa, O.V., & Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature, 31,* 463-465.

Salvi, R.J., McFadden, S.L., & Wang, J. (2000). Anatomy and physiology of the peripheral auditory system. In R. J. Roeser, M. Valente, & H. Hosford-Dunn (Eds.), *Audiology diagnosis.* New York: Thieme.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 62, 437-448.

Sanchez-Longo, L., & Forster, F. (1958). Clinical significance of impairments in sound localizations. *Neurology, 8,* 119-125.

Sand, T. (1991). BAEP amplitudes and amplitudes ratios: Relation to click polarity, rate, age, and sex. *Electroencephalography and Clinical Neurophysiology, 78*, 291-296.

Satz, P., Aschenbach, K., Pattishall, E., & Fennell, E. (1965). Order of report, ear, asymmetry, and handedness in dichotic listening. *Cortex, 1,* 203-237.

Schafer, E.W.P., & Marcus, M.M. (1973). Self-Stimulation alters human sensory brain responses. *Science, 181,* 175-177.

Schafer, E.W.P., Amochaev, A., & Russell, M. J. (1981). Knowledge of stimulus timing attenuates human cortical potentials. *Electroencephalography and Clinical Neurophysiology, 52,* 9-17.

Scharff, C., & Nottebohm, F. (1991). A compartive study of the behavioral deficits following lesions of various parts of the zebra finch song system: Implications for vocal learning. *Journal of Neuroscience, 11,* 2896-2913.

Scherg, M., & von Cramon, D. (1985). Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalography and Clinical Neurophysiology, 62*, 32-44.

Scherg, M., Hari, R., & Hämäläinen, M. (1989). Frequency-specific sources of the auditory N19-P30-P50 response detected by a multiple source analysis of evoked magnetic fields and potentials. In S. J. Williamson, M. Hoke, G. Stroink & M. Kotani (Eds.), *Advances in biomagnetism* (pp. 97-100). NY: Plenum Press.

Schow, R.L., & Chermak, G.D. (1999). Implications from factor analysis for central auditory processing disorders. *American Journal of Audiology, 8,* 137-142.

Schuller, G. (1979). Vocalization influences auditory processing in collicular neurons of the CF-FM-bat, *Rhinolophus ferrumequinum*. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *132*, 39-46.

Seki, Y., & Okanoya, K. (2008). Functional evidence for internal feedback in the songbird brain nucleus HVC. *NeuroReport, 19,* 679-682.

Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology, 23,* 695-703.

Seyfeddinipur, M., Kita, S., & Indefrey, P. (2008). How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs. *Cognition, 108,* 837-842.

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

Sharma, A., & Dorman, M.F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America, 107,* 2697-2703.

Sharma, A., Marsh, C.M., & Dorman, M.F. (2000). Relationship between N1 evoked potential morphology and the perception of voicing. *Journal of the Acoustical Society of America, 108,* 3030-3035.

Shea, S. D., & Margoliash, D. (2003). Basal forebrain cholinergic modulation of auditory activity in the zebra finch song system. *Neuron, 40*, 1213-1226.

Sheehan, J.G. (1974). Stuttering behavior: A phonetic analysis. *Journal of Communication Disorders, 7,* 193-212.

Shergill, S.S., Brammer, M.J., Amaro, E., Jr., Williams, A.C.R., Murray, R.M., & McGuire, P.K. (2004). Temporal coruse of auditory hallucinations. *British Journal of Psychiatry, 185,* 516-517.

Shergill, S.S., Brammer, M.J., Fukuda, R., Bullmore, E., Amaro, E., Jr., Murray, R.M., & McGuire, P.K. (2002). Modulation of activity in temporal cortex during generation of inner speech. *Human Brain Mapping, 16,* 219-227.

Shestakova, A., Brattico, E., Soloviev, A., Klucharev, V., & Huotilainen, M. (2004). Orderly cortical representation of vowel categories presented by multiple exemplars. *Cognitive Brain Research, 21,* 342-350.

Simson, R., Vaughn, Jr., H.G., & Ritter, W. (1977).The scalp topography of potentials in auditory and visual discrimination tasks. *Electroencephalography and Clinical Neurophysiology, 42,* 528-535.

Skinner, P.H., & Jones, H.C. (1968). Effects of signal duration and rise time on the auditory evoked potential. *Journal of Speech and Hearing Research, 11,* 301-306.

Sklare, D.A., & Lynn, G.E. (1984). Latency of the P3 event-related potential: Normative aspects and within-subject variability. *Electroencephalography and Clinical Neurophysiology, 59,* 420-424.

Smith, B., & Resnick, D. (1972). An auditory test for assessing brainstem integrity: Preliminary report. *Laryngology, 82,* 414-424.

Sommer, W., Matt, J., & Leuthold, H. (1990). Consciousness of attention and expectancy as reflected in event-related potentials and reaction times. *Journal of Experimental Psychology, 16,* 902-915.

Song, J.H., Banai, K., Russo, N.M., & Kraus, N. (2006). On the relationship between speech- and nonspeech-evoked auditory brainstem responses. *Audiology and Neuro-Otology, 11*, 233-241.

Speaks, C., & Jerger, J. (1965). Method for measurement of speech identification. *Journal of Speech and Hearing Research, 8,* 185-194.

SpectraPRO- FFT spectral Analysis system Version 3.32 User Guide.

Sperry, R.W. (1950). Neural basis of the spontaneous optokinetic response by visual inversion. *Journal of Comparative and Physiological Psychology, 43,* 482-489.

Squires, K.C., Wickens, C., Squires, N.K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science, 193,* 1142-1146.

Squires, N. K., Donchin, E., Squires, K. C., & Grossberg, S. (1977). Bisensory stimulation: Inferring decision-related processes from P300 component. *Journal of Experimental Psychology, 3,* 299-315.

Squires, N.K., Halgren, E., Wilson, C., & Crandall, P. (1983). Human endogenous limbic potentials: Cross-modality and depth/surface comparisons in epileptic subjects. In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in ERP Research: Endogenous Components* (pp. 217-232). Amsterdam, North Holland.

Squires, N.K., Squires, K.C., & Hillyard, S.A. (1975). Two varieties of long latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*, 387- 401.

Stach, B.A. (2003). Comprehensive dictionary of audiology illustrated (2nd ed.). Clifton Park, NY: Delmar Learning.

Stapells, D.R. (2002). Cortical event-related potentials to auditory stimuli. In J. Katz (Ed.), *Handbook clinical audiology* (5th ed., pp. 378-406). New York: Lippincott Williams & Wilkins.

Starr, A., & Achor, L J. (1975). Auditory brainstem responses in neurological disease. *Archives of Neurology, 32,* 761-768.

Steinschneider, M., Volkov, I.O., Noh, M.D., Garell, P.C., Howard, 3rd., M. A. (1999). Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *Journal of Neurophysiology, 8*2, 2346-2357.

Stockard, J.J., Stockard, J.E., & Sharborough, F.W. (1978). Non-pathologic factors influencing brainstem auditory evoked potentials. *American Journal of EEG Technology, 18,* 177-209.

Stockard, J.J., Hughes, J.F., & Sharborough, F.W. (1979). Visually evoked potentials to electronic pattern reversal: latency variations with gender age and technical factors. *American Journal of EEG Technology, 19,* 171-204.

Stromsta, C. (1972). Inter-aural phase disparity of stutterers and non-stutterers. *Journal of Speech and Hearing Research, 15,* 771-780.

Stuart, A., Kalinowski, J., Armson, J., Stenstrom, R., & Jones, K. (1996). Fluency effect of frequency alterations of plus/minus one-half and one-quarter octave shifts in auditory feedback of people who stutter. *Journal of Speech and Hearing Research, 39*, 396-401.

Stuart, A., Kalinowski, J., & Rastatter, M.P. (1997). Effects of monaural and binaural altered auditory feedback on stuttering frequency. *Journal of the Acoustical Society of America, 101*, 3806-3809.

Stuart, A., Kalinowski, J., Rastatter, M.P., & Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *Journal of Acoustical Society of America, 11,* 2237-2241.

Suga, N., & Jen, P.H.S. (1975). Peripheral control of acoustic signals in the auditory system of echolocating bats. *Journal of Experimental Biology, 62,* 277-311.

Suga, N., & Schlegel, P. (1972). Neural attention of responses to emitted sounds in echolocating bats. *Science, 177*, 82-84.

Suga, N., & Shimozawa, T. (1974). Site of neural attenuation of responses to self-vocalized sounds in echolocating bats. *Science, 183,* 1211-1213.

Sugg, M. J., & Polich, J. (1995). P300 from auditory stimuli: Intensity and frequency effects. *Biological Psychology, 41,* 255-269.

Sussman, E., Winkler, I., & Wang, W.J. (2003). MMN and attention: Competition for deviance detection. *Psychophysiology, 40*, 430-435.

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

Sutton, S., Roehrig, W. C., & Kramer, J. (1963). Delayed auditory feedback of speech in schizophrenics and normal's. *Annals of the New York Academy of Sciences*, *105*, 832-844.

Szymanski, M.D., Rowley, H.A., & Roberts, T.P.L. (1999). A hemispherically asymmetrical MEG response to vowels. *NeuroReport, 10,* 2481-2486.

Takaso, H., Eisner, F., Wise, R.J.S., & Scott, S. (2010). The effect of delayed auditory feedback on activity in the temporal lobe while speaking: A PET study. *Journal of Speech, Language, and Hearing Research, 53,* 226-236.

Tarkiainen, A., Hämäläinen, M., & Salmelin, R. (1997). Comparison of spherical and realistically shaped conductor models in magnetoencephalography. *International Journal of Psychophysiology, 25,* 37.

Taylor, I.K. (1966). The properties of stuttered words. *Journal of Verbal Learning and Behavior, 5,* 112-118.

Thomsen, J., Nyboe, J., Borum, P, Tos, M., & Barfoed, C. (1981). Acoustic neuromas. *Archives of Otolaryngology, 107,* 601-607.

Thornton, A.R.D. (1987). Stimulus, recording and subject factors influencing ABR diagnostic criteria. *British Journal of Audiology, 21,* 183-189.

Tiitinen, H., Sivonen, P., Alku, P., Virtanen, J., & Näätänen, R. (1999). Electromagnetic recordings reveal latency differences in speech and tone processing in humans. *Cognitive Brain Research, 8,* 355-363.

Tillman, T. (1969). Special hearing tests in otoneurological diagnosis. *Acta Otolaryngology, 89,* 25-30.

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

Tremblay, K., Friesen, L., Martin, B.A., & Wright, R. (2003). Test-retest reliability of cortical evoked potentials using naturally produced speech sounds. *Ear and Hearing, 24,* 225-232.

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear and Hearing, 22*, 79-90.

Tremblay, K., Piskosz, M., & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology, 114,* 1332-1343.

Trune, D. R., Mitchell, C., & Phillips, D. S. (1988). The relative importance of head size, gender and age on the auditory brainstem response. *Hearing Research. 32,* 165-174.

Turlough, Fitzgerald, M.J., Gruener, G., & Mtui, E. (2007). Hemispheric asymmetries. In M.J. Turlough Fitzgerald, G. Gruener, & E. Mtui (Eds). *Clinical neuroanatomay and Neuroscince* (pp.346-356). Philadelphia: Elsevier Saunders.

Uppenkamp, S., Johnsrude, I. S., Norris, D., Marslen-Wilson, W., & Patterson, R.D. (2006). Locating the initial stages of speech-sound processing in the human cortex. *NeuroImage, 31,* 1284-1296.

Vaughan, H.G., Jr., Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalography and Clinical Neurophysiology, 28*, 360-367.

Vaughan, H.G., Jr., Ritter, W., & Simon, R. (1980). Topographic analysis of auditory event-related potentials. *Progressive Brain Research, 54,* 279-285.

Ventura, M.I., Nagarajan, S.S., & Houde, J.F. (2009). Speech target modulated speaking induced suppression in auditory cortex. *BMC Neuroscience, 10:58.* doi:10.1186/1471-2202-10-58

von Holst, E. (1954). Relations between the central nervous system and the peripherial organs. *British Journal of Animal Behaviour, 2,* 89-94.

Vorobiev, V., Govoni, P., Rizzolatti, G., Matelli, M., & Luppino, G. (1998). Percolation of human medial area 6: Cytoarchitectonic evidence for three separate areas. *European Journal of Neuroscience, 10,* 2199-2203.

Wagenaar, V.A. (1969). Note on the construction of diagram-balanced Latin squares. *Psychological Bulletin, 72*, 384-386.

Webster, R.L., & Dorman, M.F. (1970). Decreases in stuttering frequency as a function of continuous and contingent forms of auditory masking. *Journal of Speech and Hearing Research, 13,* 82-86.

Webster, R.L., Schumacher, S.J., & Lubker, B.B. (1970). Changes in stuttering frequency as a function of various intervals of delayed auditory feedback. *Journal of Abnormal Psychology, 75,* 45-49.

Welch, R. B. (1978). *Perceptual modification: Adapting to altered sensory environments.* New York: Academic Press.

Wheeldon, L.R., & Levelt, W.J.M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34,* 311-334.

Whiting, K.A., Martin, B.A., & Stapells, D.R. (1998). The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da/. *Ear and Hearing, 19,* 218-231.

Wiener, F.M. & Ross, D.A. (1946). The pressure distribution in the auditory canal in a progressive sound field. *Journal of the Acoustical Society of America, 18,* 401-408.

Wiley, T.L., Oviatt, D.L., & Block, M.G. (1987). Acoustic-immittance measures in normal ears. *Journal of Speech and Hearing Research, 30,* 161-170.

Willeford, J. (1977). Assessing central auditory behavior in children: A test battery approach. In R. Keith (Ed.), *Central auditory dysfunction* (pp.43-72). New York: Grune and Stratton.

Willeford, J., & Burleigh, J. M. (1985). *Handbook of central auditory processing disorders in children.* Orlando: Grune & Stratton.

Wilson, R.H., Preece, J.P., Salamon, D.L., Sperry, J.L., & Bornstein, S.P. (1994). Effects of time compression and time compression plus reverberation on the intelligibility of the Northwestern University Auditory Test No. 6. *Journal of the American Academy of Audiology, 5,* 269-277.

Wilson, R.H., Zizz, C.A., & Sperry, J.L. (1994). Masking-level difference for spondaic words in 200-msec bursts of broadband noise. *Journal of the American Academy of Audiology, 5,* 236-242.

Wise, R. J. S., Greene, J., Buchel, C., & Scott, S. K. (1999). Brain regions involved in articulation. *Lancet, 353,* 1057-1066.

Wolpaw, J.R., & Penry, J.K. (1975). A temporal component of the auditory evoked response. *Electroencephalography and Clinical Neurophysiology, 39,* 609-620.

Wood, C.C., Allison, T., Goff, W.R., Williamson, P.D., & Spencer, D.D. (1980). On the neural origin of P300 in man. *Progress in Brain Research, 54,* 51-56.

Wood, C. C., & McCarthy, G. (1986). A possible frontal lobe contribution to scalp P300. In R. Johnson Jr., J. W. Rohrbaugh, & R. Parasuraman (Eds.), *Proceedings of the Eighth International Conference on Event-Related Potentials of the Brain* (pp. 164). Palo Alto, CA: EPIC VIII.

Wood, C.C., & Wolpaw, J.R. (1982). Scalp distribution of human auditory evoked potentials. II. evidence for overlapping sources and involvement of auditory cortex. *Electroencephalography and Clinical Neurophysiology, 54*, 25-38.

Woods, D.L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalography and Clinical Neurophysiology, 44,* (Supplement), 102-109.

Woods, D.L. Clayworth, C.C., Knight, R.T., Simpson, G.V., & Naeser, M.A. (1987). Generators of middle- and long-latency auditory evoked potentials: Implications from studies of patients with bitemporal lesions. *Electroencephalography and Clinical Neurophysiology, 68,* 132-148.

Woolley, S.M.N., & Rubel, E. (1997). Bengalese finches Lonchura striate domestic depend upon auditory feedback for the maintenance of adult song. *Journal of Neuroscience, 17,* 6380-6390.

Wunderlich, J.L., & Cone-Wesson, B.K. (2001). Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. *Journal of the Acoustical Society of America*, 109, 1526-1537.

Xu, Y., Larson, C.R., Bauer, J.J., & Hain, T.C. (2004). Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences. *Journal of the Acoustical Society of America, 116,* 1168-1178.

Yetkin, F.Z., Roland, P.S., Christensen, W.F., & Purdy, P.D. (2004). Silent functional magnetic resonance imaging (fMRI) of tonotopicity and stimulus intensity coding in human primary auditory cortex. *Laryngoscope, 114,* 512-518.

Yingling, C. D., & Hosobuchi, Y. (1984). A subcortical correlate of P300 in man. *Electroencephalography and Clinical Neurophysiology, 59,* 72-76.

Yingling, C.D., Skinner, J.E. (1977). Gating of thalamic input to cerebral cortex by nucleus reticularis larminaris. In J. E. Desmedt (Ed.), *Attention, voluntary contraction and event-related cerebral potentials. Progress in Clinical Neurophysiology* (pp. 70-96). Basel: Karger.

Yoshiura, T., Ueno, S., Iramina, K., & Masuda, K. (1995). Source localization of middle latency auditory evoked magnetic fields. *Brain Research, 703,* 139-144.

Young, L.L., Dudley, B., & Gunter, M.B. (1982). Thresholds and psychometric functions of the individual spondaic words. *Journal of Speech and Hearing Research*, *25*, 586-593.

Yue, Q., Casseday, J.H., & Covey, E. (2007). Response properties and location of neurons selective for sinusoidal frequency modulation in the inferior colliculus of the big brown bat. *Journal of Neurophysiology, 98,* 1364-1373.

Yvert, B., Crouzeix, A., Bertrand, O., Seither-Preisler, A., & Pantev C. (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cerebral Cortex, 11*, 411-423.

APPENDIX A: IRB APPROVAL LETTER

University and Medical Center Institutional Review Board
East Carolina University • Brody School of Medicine
600 Moye Boulevard • Old Health Sciences Library, Room 1L-09 • Greenville, NC 27834
Office **252-744-2914** • Fax **252-744-2284** • **www.ecu.edu/irb**
Chair and Director of Biomedical IRB: L. Wiley Nifong, MD
Chair and Director of Behavioral and Social Science IRB: Susan L. McCammon, PhD

TO: Shannon Swink, B.S. Ed., Dept of CSDI, ECU—1310 LAHN Building

FROM: UMCIRB

DATE: May 19, 2009

RE: Expedited Category Research Study

TITLE: "Auditory Monitoring During Passive Listening and Speech Production"

UMCIRB #09-0442

This research study has undergone review and approval using expedited review on 5.12.09. This research study is eligible for review under an expedited category because it is on collection of data through noninvasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves. Where medical devices are employed, they must be cleared/approved for marketing. (Studies intended to evaluate the safety and effectiveness of the medical device are not generally eligible for expedited review, including studies of cleared medical devices for new indications.) Examples: (a) physical sensors that are applied either to the surface of the body or at a distance and do not involve input of significant amounts of energy into the subject or an invasion of the subject's privacy; (b) weighing or testing sensory acuity; (c) magnetic resonance imaging; (d) electrocardiography, electroencephalography, thermography, detection of naturally occurring radioactivity, electroretinography, ultrasound, diagnostic infrared imaging, doppler blood flow, and echocardiography; (e) moderate exercise, muscular strength testing, body composition assessment, and flexibility testing where appropriate given the age, weight, and health of the individual. It is also a research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. (NOTE: Some research in this category may be exempt from the HHS regulations for the protection of human subjects. 45 CFR 46.101(b)(2) and (b)(3). This listing refers only to research that is not exempt.)
The Chairperson (or designee) deemed this **unfunded** study **no more than minimal risk** requiring a continuing review in **12 months**. Changes to this approved research may not be initiated without UMCIRB review except when necessary to eliminate an apparent immediate hazard to the participant. All unanticipated problems involving risks to participants and others must be promptly reported to the UMCIRB. The investigator must submit a continuing review/closure application to the UMCIRB prior to the date of study expiration. The investigator must adhere to all reporting requirements for this study.

The above referenced research study has been given approval for the period of **5.12.09** to **5.11.10**. The approval includes the following items:

- Internal Processing Form (dated 4.30.09)
- COI Disclosure Form (dated 4.30.09)
- Participant Demographic Information
- Informed Consent (received 5.4.09)

The Chairperson (or designee) does not have a potential for conflict of interest on this study.

The UMCIRB applies 45 CFR 46, Subparts A-D, to all research reviewed by the UMCIRB regardless of the funding source. 21 CFR 50 and 21 CFR 56 are applied to all research studies under the Food and Drug Administration regulation. The UMCIRB follows applicable International Conference on Harmonisation Good Clinical Practice guidelines.

APPENDIX B: INFORMED CONSENT- NO PARTICIPANT COMPENSATION

Title of Research Study: Auditory Monitoring During Passive Listening and Speech Production

Principal Investigator: Shannon D. Swink
Institution: East Carolina University
Address: Greenville, NC 27834
Telephone #: 252-744-6113

This consent document may contain words that you do not understand. You should ask the investigator to explain any words or information in this consent document that you do not fully understand prior to signing it.

INTRODUCTION

You have been asked to participate in a research study being conducted by Shannon D. Swink and Dr. Andrew Stuart. This research study is designed to investigate auditory monitoring in normal hearing listeners. In order to do so responses to several auditory tests will be recorded. These tests are standard audiological tests that are routinely performed in the clinical setting. Further, measurements of your brain's electrical activity ("brain waves") will be recorded by placing small surface electrodes on the scalp. If you chose to participate in this project a standard hearing assessment will be administered before any experimental procedures begin. About 2 ½-3 hours will be needed to complete the testing procedures and will be broken down into two or three sessions. All sessions will be scheduled at your convenience. All test methods used are safe, standard clinical procedures.

PLAN AND PROCEDURES

A trained audiology graduate student will perform all test procedures under the supervision of a certified clinical audiologist. There are multiple parts of this study. First, an investigator will perform a standard clinical hearing evaluation. This includes a visual inspection of the ear canal and ear drum using an horoscope (hand held light). Next you will be asked to listen to soft sounds and indicate the sound was heard through raising your hand or pressing button. To evaluate your middle ear a soft, rubber tipped probe will be gently placed in the ear canal. This probe changes the air pressure in the ear canal to measure how the eardrum moves. The procedures will take approximately 20-30 minutes.

For the next part of this study small surface electrodes will be placed on the head. The number of electrodes placed on the head will vary, ranging from three to 32. This does not require the insertion of any needle electrodes and is safe for you. You will be asked to sit in a reclining chair and either listen quietly to a series of sounds presented to both ears through headphones or vocalize sounds, which will be recorded and played back to you. During both the listen and speaking conditions electrical activity or "brain waves" at the electrodes on your head will be recorded. This procedure will take approximately 2 to 2 ½ and frequent breaks will be given during the testing procedure. All sounds employed in these experiments will be presented at a comfortable listening level.

POTENTIAL RISKS AND DISCOMFORTS

Although it is impossible to predict all possible risks or discomforts that volunteer participants may experience in any research study, the present investigators anticipate that no major risks or

discomforts will occur in the present project as standard clinical procedures are employed. Participants will not be exposed to excessive sound levels. Any risk that may be encountered would be related to mild skin irritation from skin cleansing prior to placement of the electrodes. This is usually very mild and goes away shortly. The only other form of discomfort would relate to having to sit still in the reclining chair for periods of up to one hour wearing headphones.

POTENTIAL BENEFITS

All participants will receive a free clinical hearing evaluation during the test session. Additionally, willingness to participate in this research helps East Carolina University researchers and other scientists increase their current knowledge of auditory functions in normal hearing listeners and lead to the development of new clinical procedures that will help patients in the future.

SUBJECT PRIVACY AND CONFIDENTIALITY OF RECORDS

All data collected from this study will remain confidential. Participants' names will not be used to identify the information or results in any public presentations of research findings or published research articles. Data will be coded to conceal participant identity.

COSTS OF PARTICIPATION

There will be no costs for participation.

COMPENSATION AND TREATMENT OF INJURY

The policy of East Carolina University and/or Pitt County Memorial Hospital does not provide for payment or medical care for research participants because of physical or other injury that result from this research study. Every effort will be made to make the facilities of the School of Medicine and Pitt County Memorial Hospital available for care in the event of such physical injury.

VOLUNTARY PARTICIPATION

Participating in this study is voluntary. If you decide not to be in this study after it has already started, you may stop at any time without losing benefits that you should normally receive. You may stop at any time you choose without penalty.

PERSONS TO CONTACT WITH QUESTIONS

The investigators will be available to answer any questions concerning this research, now or in the future. You may contact the investigators, Shannon D. Swink or Dr. Andrew Stuart at phone numbers 252-744-6113 or 252-744-6095. If you have questions about your rights as a research subject, you may call the Chair of the University and Medical Center Institutional Review Board at phone number 252-744-2914 (days) and/or the ECU Risk Management Office at 252-328-6858.

CONSENT TO PARTICIPATE

Title of research study: Auditory Monitoring During Passive Listening and Speech Production.

I have read all of the above information, asked questions and have received satisfactory answers in areas I did not understand. (A copy of this signed and dated consent form will be given to the person signing this form as the participant or as the participant authorized representative.)

Participant's Name **(PRINT)** **Signature** **Date** **Time**

If applicable:

Guardian's Name **(PRINT)** **Signature** **Date** **Time**

PERSON ADMINISTERING CONSENT: I have conducted the consent process and orally reviewed the contents of the consent document. I believe the participant understands the research.

Person Obtaining consent **(PRINT)** **Signature** **Date**

Principal Investigator's **(PRINT)** **Signature** **Date**

APPENDIX C: INFORMED CONSENT-PARTICIPANT COMPENSATION

Title of Research Study: Auditory Monitoring During Passive Listening and Speech Production

Principal Investigator: Shannon D. Swink
Institution: East Carolina University
Address: Greenville, NC 27834
Telephone #: 252-744-6113

This consent document may contain words that you do not understand. You should ask the investigator to explain any words or information in this consent document that you do not fully understand prior to signing it.

INTRODUCTION

You have been asked to participate in a research study being conducted by Shannon D. Swink and Dr. Andrew Stuart. This research study is designed to investigate auditory monitoring in normal hearing listeners. In order to do so responses to several auditory tests will be recorded. These tests are standard audiological tests that are routinely performed in the clinical setting. Further, measurements of your brain's electrical activity ("brain waves") will be recorded by placing small surface electrodes on the scalp. If you chose to participate in this project a standard hearing assessment will be administered before any experimental procedures begin. About 2 ½-3 hours will be needed to complete the testing procedures and will be broken down into two or three sessions. All sessions will be scheduled at your convenience. All test methods used are safe, standard clinical procedures.

PLAN AND PROCEDURES

A trained audiology graduate student will perform all test procedures under the supervision of a certified clinical audiologist. There are multiple parts of this study. First, an investigator will perform a standard clinical hearing evaluation. This includes a visual inspection of the ear canal and ear drum using an horoscope (hand held light). Next you will be asked to listen to soft sounds and indicate the sound was heard through raising your hand or pressing button. To evaluate your middle ear a soft, rubber tipped probe will be gently placed in the ear canal. This probe changes the air pressure in the ear canal to measure how the eardrum moves. The procedures will take approximately 20-30 minutes.

For the next part of this study small surface electrodes will be placed on the head. The number of electrodes placed on the head will vary, ranging from three to 32. This does not require the insertion of any needle electrodes and is safe for you. You will be asked to sit in a reclining chair and either listen quietly to a series of sounds presented to both ears through headphones or vocalize sounds that will be recorded and played back to you. During both the listen and speaking conditions electrical activity or "brain waves" at the electrodes on your head will be recorded. This procedure will take approximately 2 to 2 ½ and frequent breaks will be given during the testing procedure. All sounds employed in these experiments will be presented at a comfortable listening level.

POTENTIAL RISKS AND DISCOMFORTS

Although it is impossible to predict all possible risks or discomforts that volunteer participants may experience in any research study, the present investigators anticipate that no major risks or

discomforts will occur in the present project as standard clinical procedures are employed. Participants will not be exposed to excessive sound levels. Any risk that may be encountered would be related to mild skin irritation from skin cleansing prior to placement of the electrodes. This is usually very mild and goes away shortly. The only other form of discomfort would relate to having to sit still in the reclining chair for periods of up to one hour wearing headphones.

POTENTIAL BENEFITS

Participants will receive a stipend of $50.00 to compensate for time spent participating in this research. All participants will receive a free clinical hearing evaluation during the test session. Additionally, willingness to participate in this research helps East Carolina University researchers and other scientists increase their current knowledge of auditory functions in normal hearing listeners and lead to the development of new clinical procedures that will help patients in the future.

SUBJECT PRIVACY AND CONFIDENTIALITY OF RECORDS

All data collected from this study will remain confidential. Participants' names will not be used to identify the information or results in any public presentations of research findings or published research articles. Data will be coded to conceal participant identity.

COSTS OF PARTICIPATION

There will be no costs for participation.

COMPENSATION AND TREATMENT OF INJURY

The policy of East Carolina University and/or Pitt County Memorial Hospital does not provide for payment or medical care for research participants because of physical or other injury that result from this research study. Every effort will be made to make the facilities of the School of Medicine and Pitt County Memorial Hospital available for care in the event of such physical injury.

VOLUNTARY PARTICIPATION

Participating in this study is voluntary. If you decide not to be in this study after it has already started, you may stop at any time without losing benefits that you should normally receive. You may stop at any time you choose without penalty.

PERSONS TO CONTACT WITH QUESTIONS

The investigators will be available to answer any questions concerning this research, now or in the future. You may contact the investigators, Shannon D. Swink or Dr. Andrew Stuart at phone numbers 252-744-6113 or 252-744-6095. If you have questions about your rights as a research subject, you may call the Chair of the University and Medical Center Institutional Review Board at phone number 252-744-2914 (days) and/or the ECU Risk Management Office at 252-328-6858.

CONSENT TO PARTICIPATE

Title of research study: Auditory Monitoring During Passive Listening and Speech Production.

I have read all of the above information, asked questions and have received satisfactory answers in areas I did not understand. (A copy of this signed and dated consent form will be given to the person signing this form as the participant or as the participant authorized representative.)

Participant's Name **(PRINT)** **Signature** **Date** **Time**

If applicable:

Guardian's Name **(PRINT)** **Signature** **Date** **Time**

PERSON ADMINISTERING CONSENT: I have conducted the consent process and orally reviewed the contents of the consent document. I believe the participant understands the research.

Person Obtaining consent **(PRINT)** **Signature** **Date**

Principal Investigator's **(PRINT)** **Signature** **Date**

APPENDIX D: PARTICIPANT INTAKE QUESTIONAIRE

Auditory Monitoring During Passive Listening and Speech Production

Participant Demographic Information

Please ask an investigator any questions that may arise to ensure understanding of the demographic questions.

Age: _____

Gender: M / F (Please Circle)

Ethnicity: _____

Please Circle the Highest Educational Level Completed

High School
Post Secondary
Undergraduate Degree
Graduate Degree

Occupation: _____

Total Family Income: Please Circle

Less than $10, 000
$10, 000-$19, 999
$20, 000-$29, 999
$30, 000- $49, 999
Greater than $50, 00

Do you have a history of neurological disorders? Y or N (If Yes, please explain)

_____

Do you have a history of language disorders? Y or No (If Yes, please explain)

_____

Do you have a history of learning disabilities? Y or N (If Yes, please explain):

_____

Have you consumed alcohol or recreational drugs within the past 24 hours? Y or N

Do you take any medications that affect mental function? Y or N (If Yes, please explain):

_____

Do you have a history of hearing loss? Y or N

APPENDIX E: MINI-MENTAL STATE SCORE SHEET

Maximum	Score	
		Orientation
5	**( )**	What is the (year) (season) (date) (day) (month)?
5	**( )**	Where are we (state) (country) (town) (hospital/school) (floor)?
		Registration
3	**( )**	Name 3 objects: 1 second to say each. Then ask the subject all 3 after you have said them. Give 1 point for each correct answer. Then repeat them until he/she learns all 3. Count trials and record. Trials _____
		Attention and Calculation
5	**( )**	Serial 7's. 1 point for each correct answer. Stop after 5 answers. Alternatively spell "world" backward. (Do both and take the best score)
		Recall
3	**( )**	Ask for the 3 objects repeated above. Give 1 point for each correct answer.

Language

2	**( )**	Name a pencil and watch.
1	**( )**	Repeat the following "No ifs, ands, or buts"
3	**( )**	Follow a 3-stage command: "Take a paper in your right hand, fold it in half, and put it on the floor.
1	**( )**	Read and obey the following: CLOSE YOUR EYES
1	**( )**	Write a sentence.
1	**( )**	Copy the design shown.

Total Score: _____

Assess level of consciousness along a continuum: Alert - Drowsy

APPENDIX F: EDINBURGH HANDEDNESS INVENTORY SCORE SHEET

Please indicate your preferences in the use of hands in the following activities *by putting + in the appropriate column*. Where the preference is so strong that you would never try to use the other hand unless absolutely forced to, *put* ++. If in any case you are really indifferent *put + in both columns*.

Some of the activities require both hands. In these cases the part of the task, or object for which hand preference is wanted is indicated in brackets.

Please try to answer all the questions, and only leave a blank if you have no experience at all of the object or task.

		LEFT	RIGHT
1	Writing		
2	Drawing		
3	Throwing		
4	Scissors		
5	Toothbrush		
6	Knife (without fork)		
7	Spoon		
8	Broom (upper hand)		
9	Striking Match (match)		
10	Opening box (lid)		
	Total		

Difference	Cumulative Total	Result

APPENDIX G: SYNTHETIC TOKEN CODE AND PARAMETERS

Synthetic Male /a/ Token Code:

TIME = 000; F1=710; F2=1150; F3=2700; F0=114; AV=1

TIME + 50; F0=132; AV=72

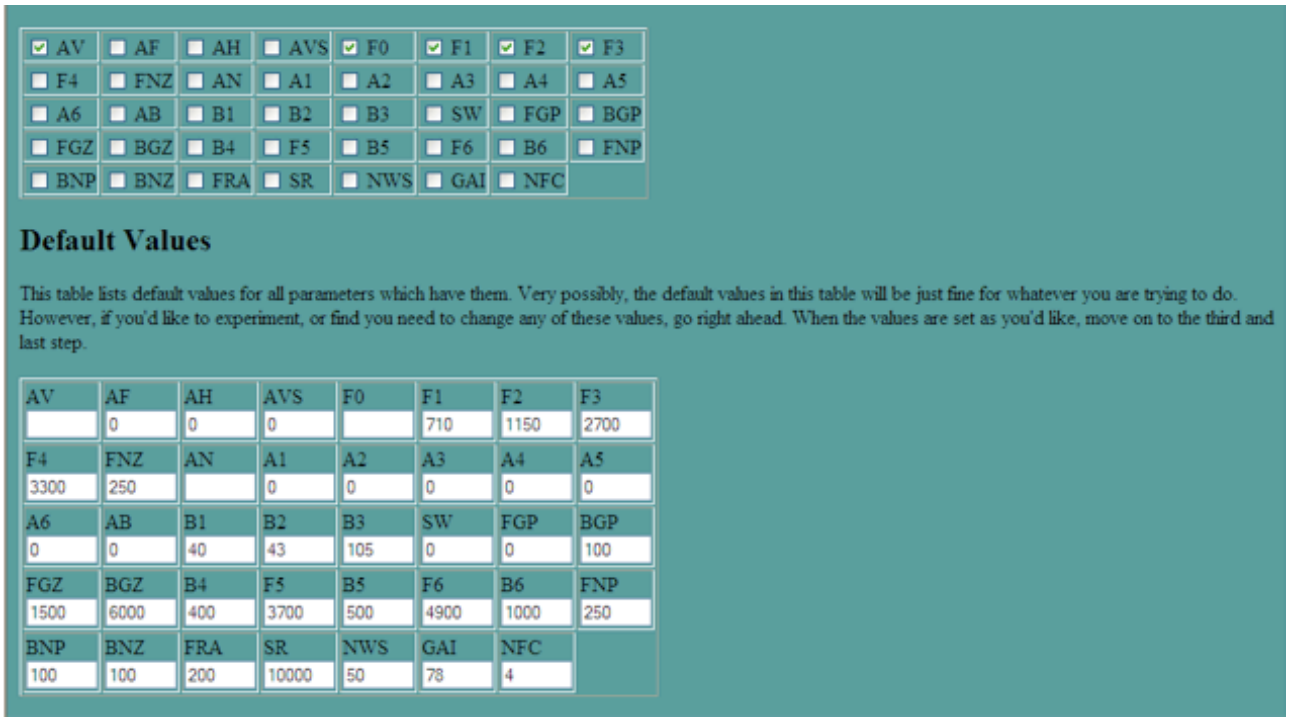
TIME = 150; F1=710; F2=1150; F3=2700; F0=132; AV=72

TIME + 200; F1=710; F2=1150; F3=2700; F0=100; AV=72

TIME + 50; AV=1

END

Synthetic Male /a/ Token Parameters:



☑ AV	☐ AF	☐ AH	☐ AVS	☑ F0	☑ F1	☑ F2	☑ F3
☐ F4	☐ FNZ	☐ AN	☐ A1	☐ A2	☐ A3	☐ A4	☐ A5
☐ A6	☐ AB	☐ B1	☐ B2	☐ B3	☐ SW	☐ FGP	☐ BGP
☐ FGZ	☐ BGZ	☐ B4	☐ F5	☐ B5	☐ F6	☐ B6	☐ FNP
☐ BNP	☐ BNZ	☐ FRA	☐ SR	☐ NWS	☐ GAI	☐ NFC	

Default Values

This table lists default values for all parameters which have them. Very possibly, the default values in this table will be just fine for whatever you are trying to do. However, if you'd like to experiment, or find you need to change any of these values, go right ahead. When the values are set as you'd like, move on to the third and last step.

AV	AF	AH	AVS	F0	F1	F2	F3
	0	0	0		710	1150	2700
F4	FNZ	AN	A1	A2	A3	A4	A5
3300	250		0	0	0	0	0
A6	AB	B1	B2	B3	SW	FGP	BGP
0	0	40	43	105	0	0	100
FGZ	BGZ	B4	F5	B5	F6	B6	FNP
1500	6000	400	3700	500	4900	1000	250
BNP	BNZ	FRA	SR	NWS	GAI	NFC	
100	100	200	10000	50	78	4	

Synthetic Female /a/ Token Code:

TIME = 000; F1=688; F2=1273; F3=2966; F0=190; AV=1

TIME + 50; F0=208; AV=72

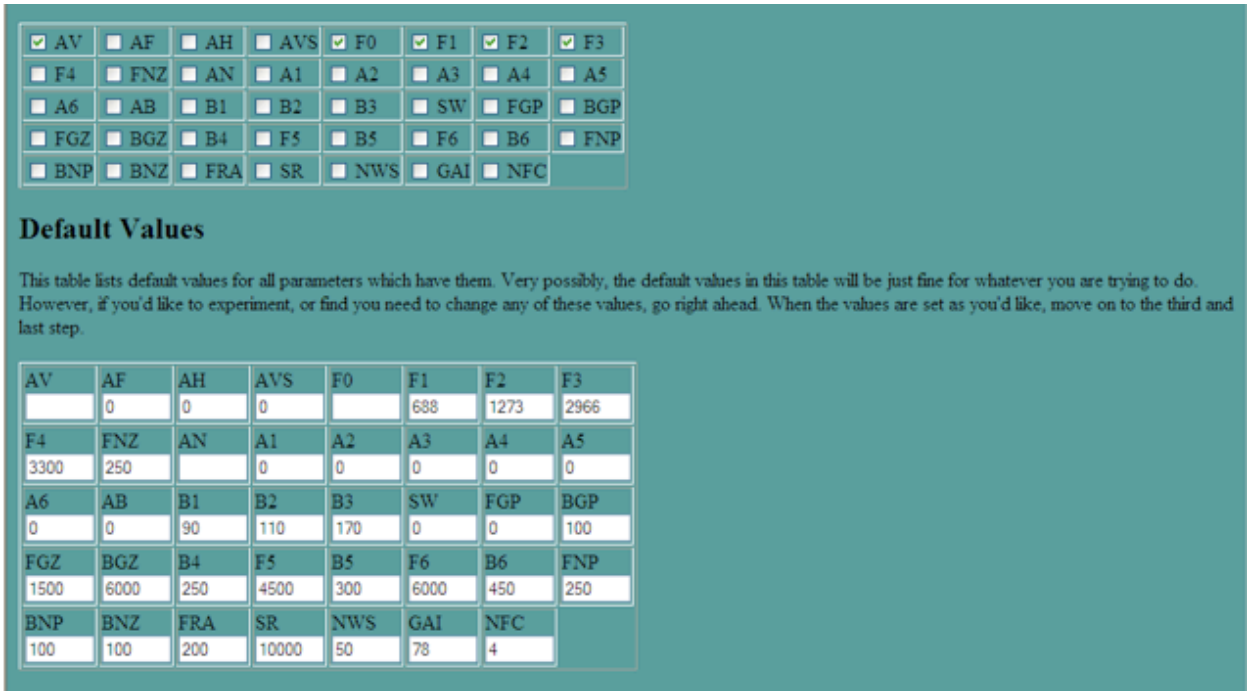
TIME + 150; F1=688; F2=1273; F3=2966; F0=208; AV=72

TIME = 200; F1=688; F2=1273; F3=2966; F0=176; AV=72

TIME + 50; AV=1

END

Synthetic Female /a/ Token Parameters:



☑ AV	☐ AF	☐ AH	☐ AVS	☑ F0	☑ F1	☑ F2	☑ F3
☐ F4	☐ FNZ	☐ AN	☐ A1	☐ A2	☐ A3	☐ A4	☐ A5
☐ A6	☐ AB	☐ B1	☐ B2	☐ B3	☐ SW	☐ FGP	☐ BGP
☐ FGZ	☐ BGZ	☐ B4	☐ F5	☐ B5	☐ F6	☐ B6	☐ FNP
☐ BNP	☐ BNZ	☐ FRA	☐ SR	☐ NWS	☐ GAI	☐ NFC	

Default Values

This table lists default values for all parameters which have them. Very possibly, the default values in this table will be just fine for whatever you are trying to do. However, if you'd like to experiment, or find you need to change any of these values, go right ahead. When the values are set as you'd like, move on to the third and last step.

AV	AF	AH	AVS	F0	F1	F2	F3
	0	0	0		688	1273	2966
F4	FNZ	AN	A1	A2	A3	A4	A5
3300	250		0	0	0	0	0
A6	AB	B1	B2	B3	SW	FGP	BGP
0	0	90	110	170	0	0	100
FGZ	BGZ	B4	F5	B5	F6	B6	FNP
1500	6000	250	4500	300	6000	450	250
BNP	BNZ	FRA	SR	NWS	GAI	NFC	
100	100	200	10000	50	78	4	

APPENDIX H: PARTICIPANT INSTRUCTIONS

Instructions for Vowel Recordings

"At a normal vocal effort and a normal speech rate please say "ah" into the microphone 10 times. These vowels are being recorded and will be used during one experimental condition.

Instructions for Passive Listening Conditions

"You will hear several different stimuli, which consists of both speech and nonspeech tokens. There will be seven different conditions. During one condition you will hear a tone burst and during the other six you will hear vowel tokens "ah" presented under various conditions. You will receive a break between each condition. During each condition I ask that you sit quietly and silently count the number of stimuli presented. Please write that number down at the end of each condition. A sheet of paper and a pencil will be given to you prior to testing. While listening to the stimulus I ask that you minimize movements especially head movements and eye blinks. Do you have any questions?"

Instructions for Speaking Conditions

"During this experiment you are asked to produce the vowel "ah" into a microphone placed a short distance from your mouth. You will hear your voice through insert earphones. This "ah" will be presented to you under various altered auditory feedback conditions. Meaning, your voice will sound normal during some conditions and will be altered during others in that it may be shifted up or delayed. Please present the vowel token until you are signaled by the examiner to stop. There will be approximately 150 presentations per trial. Please count to 2 or 3 between each vowel presentation. There will be four different conditions. You will receive a short break between each condition. Do you have any questions?

In order for your voice to be presented at a constant level there will be a training session. During this training, you are asked to hold an iPhone 3G running the application SPL meter. There is a digital display in the bottom right hand corner and gives you the dB reading. Please produce each vowel as close to 63 dB as possible. Do you have any questions?

APPENDIX I: IRB RESEARCH APPROVAL FOR CLOSURE

EAST CAROLINA UNIVERSITY
University & Medical Center Institutional Review Board Office
1L-09 Brody Medical Sciences Building• 600 Moye Boulevard • Greenville, NC 27834
Office **252-744-2914** • Fax **252-744-2284** • **www.ecu.edu/irb**

TO: Shannon Swink, BS, Department of CSDI, 1310 LAHN Building, ECU

FROM: UMCIRB

DATE: May 3, 2010

RE: Research Study Closure

TITLE: "Auditory Monitoring During Passive Listening and Speech Production"

UMCIRB #09-0442

A final review report was submitted by the investigator on 4.21.10**.** This research study has undergone expedited review for closure on 4.23.10. This **Department of CSDI** sponsored research study has been closed by the principal investigator secondary to completion of the research

It is your responsibility to ensure that you retain all research-related documents, included the informed consent forms (if applicable), for a period of no less than three years. If you have any questions or need for any reason to re-open this research study, please contact the UMCIRB Office prior to implementing any research actions.

The UMCIRB applies 45 CFR 46, Subparts A-D, to all research reviewed by the UMCIRB regardless of the funding source. 21 CFR 50 and 21 CFR 56 are applied to all research studies under the Food and Drug Administration regulation. The UMCIRB follows applicable International Conference on Harmonisation Good Clinical Practice guidelines.