

**TOWSON UNIVERSITY  
OFFICE OF GRADUATE STUDIES**

**THE EFFECT OF BACKGROUND NOISE ON SUBCORTICAL NEURAL ENCODING  
OF SPEECH IN NORMAL HEARING INDIVIDUALS**

**by**

**Laura M. Somers**

**A thesis**

**Presented to the faculty of**

**Towson University**

**in partial fulfillment**

**of the requirements for the degree**

**Doctor of Audiology**

**Department of Audiology, Speech-Language Pathology and Deaf Studies**

**Towson University**

**Towson, Maryland 21252**

**May 2016**

**TOWSON UNIVERSITY  
OFFICE OF GRADUATE STUDIES**

**THESIS APPROVAL PAGE**

This is to certify that the thesis prepared by Laura Somers B.A., Au.D. Candidate,  
entitled The Effect of Background Noise on Subcortical Neural Encoding of Speech in Normal  
Hearing Individuals has been approved by the thesis committee as satisfactorily completing the  
thesis requirements for the degree Doctor of Audiology (Au.D.).

A. Sarada  
Saradha Ananthakrishnan, Ph.D.  
Chair, Thesis Committee

5-15-2015  
Date

Peggy Korczak  
Peggy Korczak, Ph.D.  
Committee Member

5-15-2015  
Date

Stephen Pallett  
Stephen Pallett, Au.D.  
Committee Member

5/13/2015  
Date

Janet V. Delany  
Janet Delany  
Dean of Graduate Studies

5-15-15  
Date

## **ACKNOWLEDGEMENTS**

I would like to thank my Thesis committee, Dr. Saradha Ananthakrishnan, Dr. Peggy Korczak, and Dr. Steve Pallet for their continued support. I am extremely grateful for your assistance and comments throughout the duration of my Thesis. The completion of my Thesis would not have been possible without you. Thank you to Dr. Evitts for sharing your knowledge about speech acoustics and letting us borrow your recording equipment. I would also like to thank all my friends and family for their un-ending encouragement every step of the way.

## TABLE OF CONTENTS

THESIS APPROVAL.....	ii
ACKNOWLEDGEMENTS.....	iii
TABLE OF CONTENTS.....	iv
LIST OF FIGURES.....	vii
LIST OF TABLES.....	ix
ABSTRACT.....	x
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: LITERATURE REVIEW.....	2
Overview of Auditory Evoked Potentials.....	2
Basic Speech Acoustics.....	5
Auditory Neural Encoding.....	8
<i>Envelope and Temporal Fine Structure</i> .....	9
<i>Phase locking</i> .....	10
History of the Frequency Following Response (FFR).....	11
<i>Animal Studies</i> .....	12
<i>Human Studies</i> .....	15
Neural Generators of the FFR.....	17
<i>Animal Studies</i> .....	17
<i>Human Studies</i> .....	20
Basics of the FFR .....	22
Stimulus Parameters.....	29
<i>Stimulus type</i> .....	29
<i>Stimulus Intensity</i> .....	34
<i>Stimulus Rate</i> .....	36

<i>Stimulus Polarity</i> .....	38
Recording Parameters.....	39
<i>Band-Pass Filter</i> .....	39
<i>Length of the Analysis Window</i> .....	41
<i>Number of Sweeps</i> .....	42
<i>Electrode Montage</i> .....	43
<i>Sampling Rate</i> .....	45
<i>Artifact Rejection</i> .....	47
Subject Parameters.....	49
<i>Subject State</i> .....	49
<i>Subject Age</i> .....	51
Background Noise and Speech Acoustics.....	52
<i>Vowel Perception in Noise</i> .....	56
Background Noise and Neural Encoding.....	61
CHAPTER 3: Methodology.....	64
<i>Participants</i> .....	64
<i>Stimuli</i> .....	64
<i>Data Collection</i> .....	65
Data Analysis.....	66
CHAPTER 4: RESULTS.....	68
<i>Grand Average FFR Temporal waveforms (<math>FFR_{ENV}</math>)</i> .....	68
<i>Spectral Analysis: Fast Fourier Transform (FFT) of the <math>FFR_{ENV}</math></i> .....	70
<i>Quantitative Spectral Analysis (<math>FFR_{TFS}</math>)</i> .....	72
<i>Grand Average FFR Temporal waveforms (<math>FFR_{TFS}</math>)</i> .....	75
<i>Spectral Analysis: Fast Fourier Transform (FFT) of the <math>FFR_{TFS}</math></i> .....	77
<i>Quantitative Spectral Analysis (<math>FFR_{TFS}</math>)</i> .....	79
<i>Summary of Results</i> .....	83

CHAPTER 5: DISCUSSION.....	84
<i>Overview of Neural Encoding at the Basilar Membrane</i> .....	84
<i>Effects of Background Noise on FFR<sub>ENV</sub></i> .....	87
<i>Effects of Background Noise on FFR<sub>TFS</sub></i> .....	90
<i>Clinical Implications</i> .....	93
<i>Limitations</i> .....	93
<i>Future Studies</i> .....	94
APPENDICES.....	95
<i>Appendix A</i> .....	95
<i>Appendix B</i> .....	96
References.....	97
CURRICULUM VITAE.....	105

## LIST OF FIGURES

Figure 1: Narrowband spectrogram of the vowel /u/.....	7
Figure 2: Wideband spectrogram of the vowel /u/.....	7
Figure 3: Varying Envelope and TFS of a complex waveform.....	9
Figure 4: Neural firing pattern.....	10
Figure 5: The Cochlear Microphonic and FFR recorded at electrodes when the cochlear nucleus is cooled.....	14
Figure 6: Evoked Electrical Response and FFR at the SOC.....	15
Figure 7: Ipsilateral and contralateral pathways of the FFR neural generators.....	19
Figure 8: The overall spectral characteristics of the stimulus (a) and the envelope of the response (b) represented in a temporal waveform (Akhoun et al., 2008a).....	24
Figure 9: Waveform recorded to the condensation polarity of the stimulus.....	25
Figure 10: Waveform recorded to the rarefaction polarity of the stimulus.....	25
Figure 11: Condensation + Rarefaction waveform.....	25
Figure 12: Condensation – Rarefaction waveform.....	26
Figure 13: Summed envelope FFR and FFT.....	27
Figure 14: Difference waveforms representing the TFS of the stimulus and FFT.....	28
Figure 15: Types of complex and simple stimuli.....	30
Figure 16: F0 changing over time for four lexical tones (Krishnan et al., 2005).....	31
Figure 17: The FFR mimicking the period the stimulus /da/ (Cunningham et al., 2001).....	34
Figure 18: Comparison between SOA and ISI.....	36

Figure 19: Pre-stimulus part of the analysis window, the duration of the stimulus, and the post-stimulus part of the analysis window (Gong et al., 2013).....	42
Figure 20: Electrode Placement (International 10-20 system).....	45
Figure 21: Example of sampling rate (Atcherson & Stoody, 2012).....	46
Figure 22: FFTs of vowel /e/ in quiet, in the presence of white noise, and in the presence of multi talker babble at 0 dB SNR (Assmann & Summerfield, 2014).....	54
Figure 23: Grand average temporal waveforms for the summed envelope FFR ( $FFR_{ENV}$ ) for the four background noise conditions.....	69
Figure 24: $FFR_{ENV}$ grand average FFT analysis for each of the four background noise conditions.....	71
Figure 25: Average F0 amplitude obtained using FFT analysis of the summed average temporal waveform in the clean, + 5 SNR, 0 SNR, and -5 SNR listening conditions.....	74
Figure 26: Subtracted grand average temporal waveforms for four listening conditions.....	76
Figure 27: Subtracted grand mean FFT analysis for four listening conditions.....	78
Figure 28: Mean SNR measured for the four listening conditions.....	82
Figure 29: Overview of Neural Encoding at the Basilar membrane.....	86



## LIST OF TABLES

Table 1: Sampling Rates for FFR Studies.....	47
Table 2: F1 and F2 formant frequency regions for vowel like stimuli /æ/, /ε/, /ɪ/, and /i/.....	58
Table 3: Mean latency FFR values for children with HINT scores in the top and bottom 50 <sup>th</sup> percentile and the mean latencies for the top and bottom readers.....	63
Table 4: Sample data table: Amplitude (uV) of the response at 120 Hz determined by the FFT analysis.....	67
Table 5: Individual amplitude measurements at the F0, Mean F0 values, and Standard Deviation.....	73
Table 6: Individual Derived SNR measurements at the F1 and their Mean and Standard Deviation Values.....	81

## ABSTRACT

The purpose of the current research is to investigate the effects of background noise on subcortical neural encoding of speech in normal hearing individuals. It was proposed that increasing levels of background noise would degrade the Frequency Following Response (FFR). Six participants (age 23- 24 years) participated in the current study. The Frequency Following Response was recorded on all six subjects in four different conditions: clean, 0 dB SNR, + 5 dB SNR, and -5 dB SNR. Grand average temporal waveforms were obtained across the 6 subjects for each listening condition. MATLAB software was used to analyze the temporal waveforms in the frequency domain. An FFT analysis was used to break down the response into the constituent frequencies. The magnitude of energy in the FFR spectrum at the fundamental frequency and first formant were measured in each test condition for each individual participant. Mean FFT values and standard deviation were also calculated. Results showed that the F0 magnitude and derived SNR for F1 decreased with increasing levels of background noise. Overall, it was concluded that the FFR was most degraded in the most severe background noise listening condition.

## **CHAPTER 1: INTRODUCTION**

Noisy environments can be difficult to hear in, especially for persons with a hearing impairment. Often times the number one complaint of an individual with a hearing loss is that they have difficulty understanding speech in background noise. Hearing aids are designed to help suppress the background noise in order to help improve speech clarity. Despite the sophisticated algorithms in the hearing aids to help alleviate the difficulty of understanding speech in background noise, individuals with hearing loss continue to experience problems. To better understand why individuals with a hearing loss struggle with speech perception in background noise, we must first investigate the effects of background noise on a normally functioning auditory system.

The Frequency Following Response (FFR) is an electrophysiological response that can be used to assess the effects of background noise on auditory neural encoding of a speech stimulus. In particular, the FFR allows the investigator to view how well the auditory system is able to encode specific acoustic features of speech such as the F0, represented by the slow varying envelope (E), and the higher frequency harmonics, represented by the temporal fine structure (TFS) of a speech stimulus.

Leek and Summers (1996) reported that those individuals with hearing loss have a damaged underlying auditory physiology that essentially allows more noise to enter the auditory system. Once the noise enters the auditory system, this disrupts neural encoding of both the E and the TFS of the speech signal. Several researchers have reported that once the level of the background noise reaches a certain threshold, brainstem neural encoding is degraded (Li & Jeng, 2011). However, most of the current research has focused on behavioral responses, such as

speech perception in noise (Anderson, Skoe, Chandrasekaran, & Kraus, 2010; Hedrick & Nabalek, 2004; Nabalek & Dagenais, 1986). Further research is needed to truly understand the effects of background noise on auditory neural encoding. Therefore, the current study will investigate the effects of background noise on the FFR in normal hearing participants.

Before the audiologist is able to record and understand the FFR, there are various topics that must be discussed beforehand. A general background will be provided on concepts related to basic speech acoustics, neural encoding at different levels of the auditory system, the history of the FFR, the information obtained using the FFR, and the effects of background noise on the FFR. In addition to these areas, the literature review will discuss the appropriate recording, stimulus, and subject parameters for recording the FFR.

## **CHAPTER 2: LITERATURE REVIEW**

### **General Description of Auditory Evoked Potentials**

The human sensory organs are responsible for sending important information to the brain via a pathway of interconnected neurons (Hall, 2007). While there is always constant electrical activity in the brain, the presentation of any stimulus that stimulates the senses can cause a change in this ongoing electrical activity. Evoked potentials “cause fluctuations in voltage [and can be] recorded from scalp electrodes in humans” (Kraus & Nicol, 2005). An auditory evoked potential (AEP) is a change in the electrical or neuronal activity that occurs when acoustic sounds are presented to the ear.

AEPs are either classified as sensory AEPs or processing contingent potentials (PCPs). The sensory auditory evoked potential is “obligatory”, i.e. the sensory AEP is completely dependent on the physical characteristics of the stimulus and this AEP will reflect changes in the stimulus. PCPs are changes in the neuronal activity due to higher brain processing and are affected by the “arousal level and attention” of the individual (Stapells, 2009). Both the sensory AEPs and PCPs can be classified further based on the anatomical generators of the response, the latency of the response relative to the stimulus, and the overall relationship between the response and the stimulus.

AEPs can occur at any location along the auditory pathway from the cochlea to the auditory cortex (Hall, 2007). The auditory pathway begins at the cochlea, then travel along the auditory branch of the vestibular-cochlear nerve to the brainstem and cortex. The brainstem is composed of the medulla, pons, and midbrain. The cochlear nucleus (CN) and the superior olivary complex (SOC) are two structures at the level of pons that connect with the inferior colliculi of the midbrain via the lateral lemniscus (Biacabe, Chevallier, Avan, & Bonfils, 2001; Chiappa, 1997). Sensory AEPs are recorded at the subcortical level or early cortical level (Stapells, 2009). PCPs on the other hand are generated at the level of the cortex (Stapells, 2009). AEPs recorded at the subcortical and cortical level can provide information about the function of the normal and damaged auditory pathways (Hall, 2007).

The anatomical source of the AEP can be predicted by analyzing the latency or timing of the response after the onset of the stimulus (Stapells, 2009). AEPs are divided into five groups based on the latency of the response: first (0- 5 ms), fast (2-20 ms), middle (10 – 100 ms), slow (50- 300 ms), and late (150 – 1000 ms). Sensory AEPs include the first, fast, middle, and slow cortical response (waves P<sub>1</sub>, N<sub>1</sub>, and P<sub>2</sub>), whereas the late cortical responses, such as waves P3b,

are PCPs. AEPs that occur within the first 20 ms of stimulus onset are typically generated early on in the auditory pathway from anatomical generators such as the cochlea, 8<sup>th</sup> nerve, and/ or brainstem (Hall, 2007; Stapells, 2009). For example, the Auditory Brainstem Response (ABR) is a short latency or fast AEP occurring within 10 ms after the onset of the stimulus (Hall, 2007). Other evoked potentials generated from higher levels within the auditory system, e.g. closer to the cortex, exhibit longer latencies such as the slow and late AEPs (Hall, 2007; Stapells, 2009).

AEPs can also be classified based on the relationship between the stimulus and the response. The three different classifications based on this relationship are transient, steady state, and sustained (Stapells, 2009). Transient AEPs, such as the ABR, occur directly after the onset of the stimulus (Kraus & Nicol, 2005). Transient responses are defined as “fast response peaks lasting fractions of a millisecond, evoked by a non- sustained stimulus” (Skoe & Kraus, 2010). Examples of transient stimuli are clicks or tones, and the onset and offset of speech (Skoe & Kraus, 2010). A robust transient response is elicited with stimuli such as clicks because the click stimulates many neurons at once that fire synchronously (Skoe & Kraus, 2010). A sustained response is a neural response that is “time-locked” or “phase-locked” to the period of the stimulus (Skoe & Kraus, 2010). An example of a stimulus that elicits a sustained response is speech (Skoe & Kraus, 2010). The steady state response is an electrical response to a regularly-repeating stimulus that lasts throughout the duration of the stimulus.

The focus of the current study is an AEP known as the Frequency Following Response (FFR). Based on the four different classification schemes described above the FFR can be described as a fast (2- 20 ms), sustained, sensory AEP occurring at the level of the brainstem. The remainder of this literature review will focus on the FFR.

## **Basic Speech Acoustics**

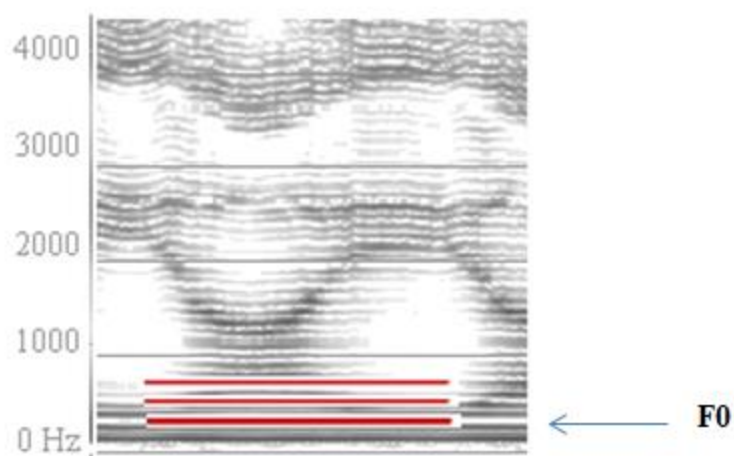
To effectively administer and interpret the FFR, audiologists and hearing scientists must have a basic understanding of speech acoustics. Below is a brief description of several key acoustic parameters present in a speech signal.

Speech is a complex signal which consists of different vowels, consonants, and phonemes (Kraus & Nicol, 2005). The source-filter theory describes how these speech sounds are produced by a source (the vocal chords) and then shaped or filtered by the vocal tract. During the production of a speech stimulus, the vocal chords (source) vibrate at a certain rate/sec, which is defined as the fundamental frequency and is labeled as F0 (Borden & Harris, 1984). The speech signal also consists of harmonic and formant frequencies. Harmonics are defined as “frequency components of a complex waveform that are whole number multiples of its fundamental frequency” (Emanuel & Letowski, 2009). For example, a stimulus with a fundamental frequency of 110 Hz will have harmonic frequencies present at 220 Hz, 440 Hz, 880 Hz, etc. The fundamental frequency can be referred to as the first harmonic of the complex waveform (Emanuel & Letowski, 2009). Formants are peaks of energy that occur at certain frequencies in a speech stimulus, such as a vowel, and are labeled F<sub>1</sub>, F<sub>2</sub>, F<sub>3</sub>, and F<sub>4</sub>. The formants of the speech stimulus are determined by the “filter” or the shape of the resonators such as the lips, tongue, and mouth (Aiken & Picton, 2008). For instance, the vowel /u/ is produced differently than the vowel /i/. The tongue placement during the production of the vowel /u/ is high and further back in the oral cavity, whereas the tongue is high and in front of the oral cavity for /i/. The manipulation of the oral “filters” causes more energy to be present at certain frequencies than at others and thus creates these formant frequencies. These formant frequencies help in the identification of the vowel (Kraus & Nicol, 2005).

The fundamental and harmonic frequencies along with the formants or peaks of energy in a complex speech waveform can be visually represented using a spectrogram. Spectrograms represent these harmonic and formant frequencies as a function of time. The spectrogram can be manipulated to help better define either the frequencies or the formants of the signal. A narrow band spectrogram is manipulated to emphasize the different frequencies (fundamental and harmonics) in the stimulus. A wide band spectrogram is manipulated to emphasize the formants or peaks in energy.

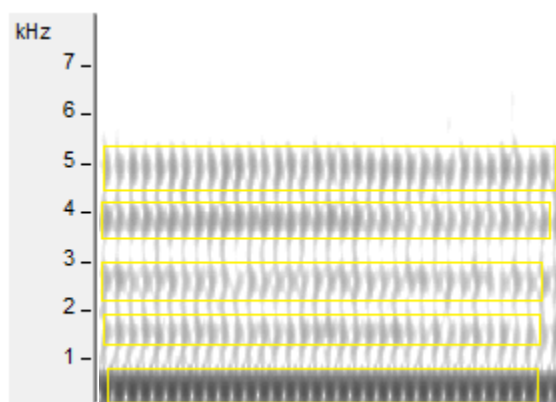
The following is an example of the information that is present in a spectrogram. The following two figures (Figure 1 and 2) highlight the difference between a narrowband vs. a wideband spectrogram. In this example a speech stimulus /u/ with a fundamental frequency of 120 Hz was recorded and analyzed using a computer software program designed to analyze speech. A narrowband spectrogram of the stimulus was created and is presented below. Note that the bottom red line indicates the fundamental frequency and the other horizontal striations represent the harmonic frequencies. The first two harmonics are represented by the bolded red lines. The harmonic frequencies of the stimulus /u/ are whole number multiples of the fundamental frequency (i.e. 240 Hz, 360 Hz).





*Figure 1.* Narrowband spectrogram of the vowel /u/

A wide band spectrogram was also created and is represented below. This spectrogram better represents the formant frequencies of the vowel /u/ due to its more precise temporal resolution. As mentioned before the formant frequencies are bands of energy that are present at particular frequencies due to the manner of articulation and shape of the oral cavity during the production of the vowel /u/ vowel. Note that the formant frequencies are the dark banded regions as represented by the yellow boxes.



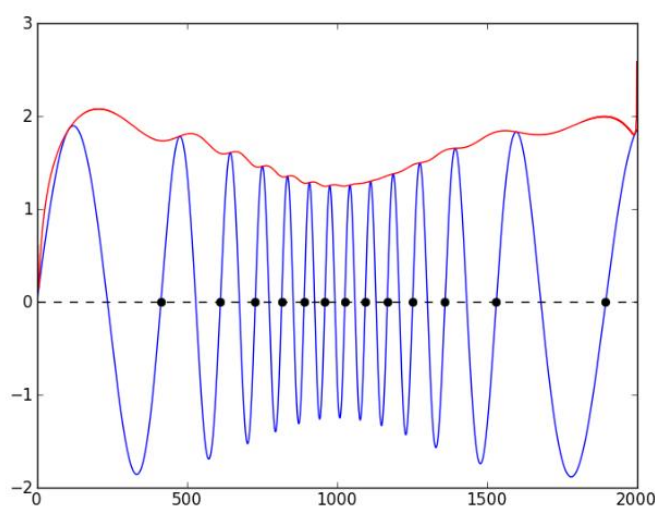
*Figure 2.* Wideband spectrogram of the vowel /u/

An understanding of the fundamental frequency, harmonic frequencies, and formant frequencies is important because the FFR is capable of reflecting both the fundamental frequency as well as the formant structure contained in the speech signal. Before the FFR can be discussed in further detail a better understanding of the underlying physiologic process for speech encoding is necessary.

### **Neural Encoding**

Neural encoding begins in the peripheral auditory system, specifically at the level of the basilar membrane. The encoding of a sound stimulus, whether simple or complex, begins with “band pass filters” along the basilar membrane (Moore, 2008). The band pass filters consist of a lower and upper frequency cut-off, which allow a certain frequency or group of frequencies within these cut-offs to pass through. For instance, a band pass filter with a lower frequency of 500 Hz and an upper frequency of 600 Hz will encode a 525 Hz stimulus. The basilar membrane is composed of a series of such overlapping band pass filters. These auditory filters follow the tonotopic mapping of the cochlea, with filters at the basal end centered around higher frequencies, while filters at the apex are centered around lower frequencies. Further, filters located at the base of the cochlea are wider in range than the filters located at the apex of the cochlea. For instance the filter at the basal end of the cochlea may extend from 3000 Hz to 6050 Hz, therefore allowing a broad range of frequencies to be processed and encoded through that filter. On the other hand, the filter at the apical end only ranges from 100 Hz to 200 Hz thus allowing a small range of frequencies through. A single low frequency pure tone passing through these band pass filters at the apical end of the cochlea will result in a single pure tone that exits the cochlea and is further encoded by the 8<sup>th</sup> nerve. For example, a 150 Hz pure tone stimulus will pass through the 100 Hz and 200 Hz band pass filter and exit as the same 150 Hz pure tone.

In contrast, the multiple high frequencies passing through the broad band pass filters at the basal end of the cochlea will overlap and create a complex waveform. This complex waveform consists of a slowly varying envelope superimposed on a rapidly oscillating temporal fine structure (Moore, 2008). The figure below generally depicts the relationship between the envelope and the TFS of a complex waveform. As seen in this figure the red line is the envelope and the blue is the TFS.

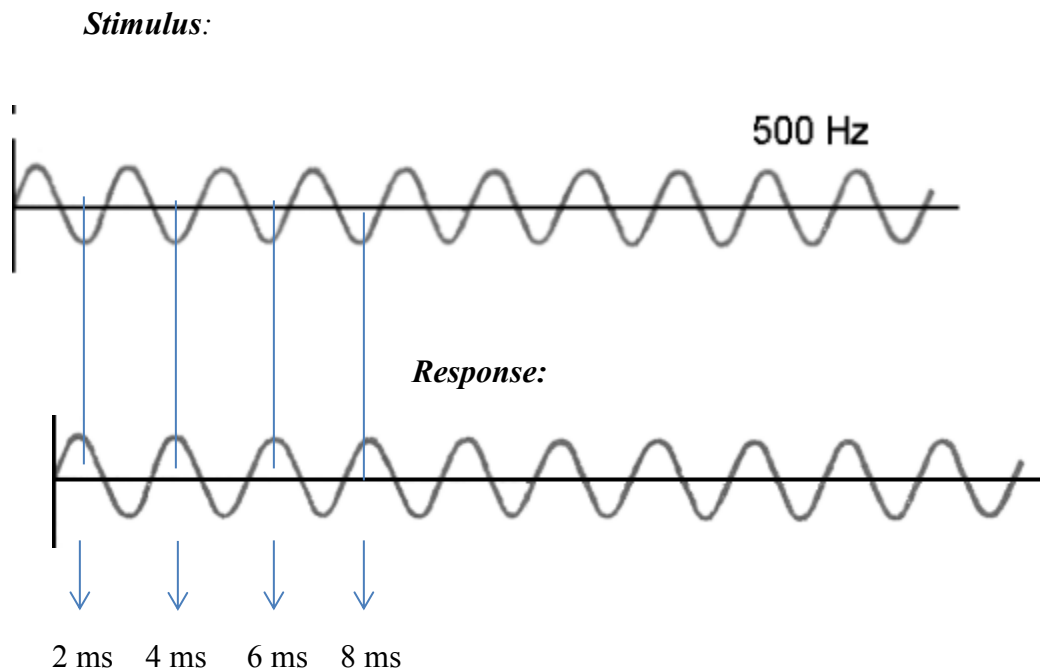


*Figure 3. Varying Envelope and TFS of a complex waveform*

These two spectral properties, the envelope and the temporal fine structure (TFS), are further encoded at the level of the 8<sup>th</sup> nerve and brainstem where a phenomenon called “phase locking” takes places. The envelope of the waveform is related to the fundamental frequency of the stimulus (F0) and the TFS is related to the formants of the stimulus. The neurons within the brainstem exhibit these phase locking properties and will be discussed further in the next section.

### ***Phase Locking Properties of the 8<sup>th</sup> nerve***

Phase locking is an important physiological phenomenon underlying the FFR. A phase locked response is defined as “neuronal discharges that occur at whole number multiples of the period of the stimulus” (Roullier, Ribaupierre, & Ribaupierre, 1979). Recall that the period of the stimulus is the reciprocal of the fundamental frequency. For example if the frequency of the stimulus is 500 Hz, the period of the stimulus ( $1000 \text{ ms} / 500 \text{ Hz}$ ) is 2 ms. Therefore, the neuronal discharges occur at 2 ms intervals (2 ms, 4 ms, 6 ms, 8 ms, etc...). Below is a figure that visually depicts a 500 Hz stimulus and how the response “follows” the stimulus at 2 ms, 4ms, 6 ms, and 8 ms, etc. The neural firing pattern is indicated by the solid lines.



*Figure 4. Neural firing pattern*

Both envelope and TFS information in a speech signal are encoded via neural phase-locking. The FFR is considered a “phase locked” response since the brainstem neurons generating this response fire at intervals corresponding to the fundamental and formant frequencies of the stimulus. Hence the FFR represents neural phase-locking to both envelope and TFS information at the rostral brainstem level. With this basic understanding of neural encoding and phase locking, we can proceed to a more detailed discussion of the FFR as a tool for evaluating neural encoding of speech.

### **History of the FFR**

The FFR is a scalp recorded sustained evoked potential that reflects neural phase-locking from the rostral brainstem. The research behind the discovery of this neural response and the neural origins of the FFR are discussed in greater detail below

A long standing debate in the AEP literature regarding the FFR was whether or not the response was a true neural response or the result of a “non-neural mechanism” (Marsh, Worden, & Smith, 1970). Marsh and colleagues (1970) reported that possible non-neural origins of the FFR included the cochlear microphonic (CM) as well as stimulus artifacts created at the electrode sites. In order to explore the origin of the FFR, investigators have conducted a series of animal and human experiments. The results of the animal studies will be reported first, followed by reports from human studies.

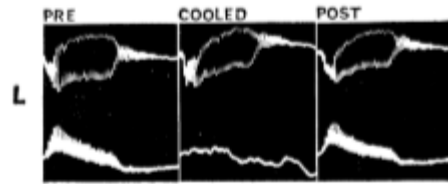
### ***Results from animal recordings***

In 1930, Wever and Bray conducted an experiment to determine whether the response to an auditory stimulus was of neural or cochlear origin. In this study, these investigators conducted near-field recordings on cats with needle electrodes located on the eighth nerve as well as on other structures in the brain. Wever and Bray (1930) initially believed that the cochlea was the origin of this near-field response. However, there were several pieces of evidence that supported that the amplitude or magnitude of this electrophysiologic response decreased after the cochlea was destroyed, thus suggesting that the response was not of cochlear origin. Specifically, Wever and Bray (1930) reported that the neural responses were localized to certain areas of the brain, such as the medulla and the brainstem. Wever and Bray (1930) also speculated that the distance between the electrode placement in the brain (near-field recording) and the cochlea would most likely prohibit a cochlear mechanism from contributing to the response. Lastly, the responses recorded were “localized to specific nervous tracts”, which would not be the case if the response was originating from the cochlea (Wever & Bray, 1930).

In 1970, Marsh and colleagues conducted three experiments on animals to determine if the FFR was a true neural response or a stimulus artifact. In the first experiment, Marsh and colleagues recorded the CM and the neural response to a 730 Hz stimulus with needle electrodes placed at the round window for the CM and at the cochlear nucleus for the FFR. Responses were recorded 5 minutes prior to and directly following the sectioning of the eighth nerve. These investigators reported that the FFR at the cochlear nucleus disappeared as soon as the eighth nerve was severed and remained absent five minutes following the sectioning. In contrast, the cochlear microphonic recorded at the round window was present following the eighth nerve

section. These results provide clear evidence that the FFR is a true neural response and is dependent on an intact neural pathway (Marsh et al., 1970).

In the second experiment, Marsh and colleagues sought to confirm the neural origin of the FFR. In this second experiment, responses were recorded again using electrodes located at the round window and at the cochlear nucleus. A cryoprobe was also inserted into the cochlear nucleus during the recordings. The purpose of the cryoprobe was to cool the cochlear nucleus and essentially act as a “reversible blockage” of the FFR. Marsh and colleagues reported that as the temperature decreased by approximately 6 degrees (Celsius) below its normal temperature, the FFR disappeared. The FFR re-appeared when the temperature returned to normal. In contrast, the CM response was present throughout the cooling process (Marsh et al., 1970). As seen below, Figure 4 depicts the CM (top response) and the FFR (bottom response) in each panel. The left panel depicts the test condition before the cochlear nucleus was cooled. The mid panel shows the response during cooling and the right panel depicts the post cooling period. The “L” denotes that the stimulus was delivered to the left ear. Note that the CM response remains unchanged throughout the three conditions (pre, cooled, and post). However, the FFR was not present during the cooling, but is robust before and after the cooling of the CN. Collectively, these results of these two experiments further support that the FFR “depends upon the functional integrity of the auditory pathway” (Marsh et al., 1970).



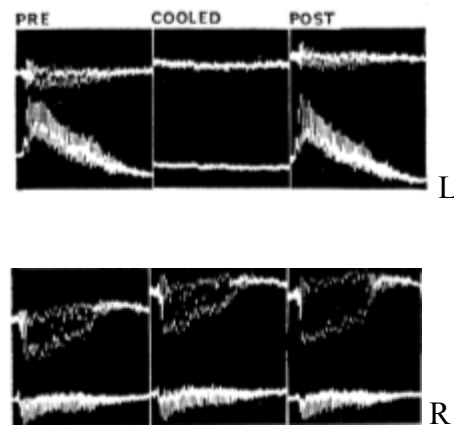
*Figure 5.* The Cochlear Microphonic and FFR recorded at electrodes when the cochlear nucleus is cooled. Electrodes placed near the round window and cochlear nucleus to record the CM (top response) and FFR (bottom response) respectively. The FFR is not present during cooling of the cochlear nucleus (middle response).

In order to further investigate the neural origin of the FFR, Marsh et al. (1970) placed additional needle electrodes at the level of superior olivary complex (SOC), which is higher in the brainstem in comparison to the CN, in order to determine if the FFR could be recorded at the SOC during cooling of the CN. In this third experiment, the FFR was recorded at the right and left superior olive nuclei as the left cochlear nucleus was cooled. A 900 Hz stimulus was presented to the left ear, the right ear, and binaurally. Figure 5 was extracted from Marsh et al. (1970) to visually show the effects of the cooled left cochlear nucleus on recordings obtained from the SOC. The evoked electrical responses (top) and FFR (bottom) at the SOC are pictured in the figure below with the stimulus presented to the left and right ear (top and bottom response of Figure). As seen in the top row, Marsh et al. (1970) reported that the FFR was abolished at the SOC when the left cochlear nucleus was cooled and the stimulus was presented ipsilaterally (left ear). However, as shown in the second row of the bottom response, the SOC and FFR recorded at this anatomical site is unaffected when the left cochlear nucleus was cooled and the stimulus was delivered to the contralateral ear (right ear). Thus, cooling of the CN prohibited the neural response from being recorded at the ipsilateral SOC only. Therefore, Marsh et al. (1970)



concluded that the FFR is a neural response which originates partly due to contributors from the CN.

Collectively, the findings from these animal studies confirmed that the origin of the CM was in the cochlea, while the FFR was of neural origin. One neural structure that contributed to the FFR in animals was the cochlear nucleus.



*Figure 6.* Evoked Electrical Response and FFR at the Superior Olivary Complex. The top response in both figures represents the evoked electrical potential and the bottom response represents the FFR. “L” denotes that the stimulus was delivered to the left ear. “R” denotes that the stimulus delivered to the right ear.

### ***Results from Human Recordings***

Approximately 5 years later, Marsh and colleagues expanded their research on the origin of the FFR to include human subjects. They then compared the FFR recordings obtained from humans to those of cats. Smith, Marsh, and Brown (1975) obtained far-field recordings of FFR in humans with electrodes placed at various locations on the scalp. The FFR obtained from humans

in this experiment using the far-field recordings were then compared to the responses in cats obtained using near-field recordings in earlier studies. The needle electrodes placed in the cats were located at several areas in the brainstem including the CN, SOC, and inferior colliculus (IC). The stimulus that was used during the near field recordings ranged from 500 Hz to 1000 Hz. In contrast, during the far field recordings in humans, only a 500 Hz stimulus was presented at an intensity level of 60 to 70 dB suprathreshold (Smith et al., 1975). The latency of the FFR was then measured in both populations (latency is measured as the amount of time, in milliseconds, it takes for the response to occur following the onset of the stimulus). The latency of the FFR recorded at the IC in cats was similar to the latency of the FFR recorded using surface electrodes on humans. Therefore, Smith et al. (1975) concluded that the human scalp-recorded FFR reflected neural activity occurring within the brainstem. Furthermore, these results confirmed the results from earlier animal studies, which suggested that the FFR was of neural origin. This study also further investigated specific regions within the auditory brainstem region that contribute to the FFR and therefore the results of this study are discussed in more detail in the section on Neural Generators.

More recent review articles of the FFR have discussed key characteristics that help to distinguish the FFR from the CM (Skoe & Kraus, 2010). These investigators reported the CM and FFR both “mimic the temporal waveform of the acoustic stimulus”; however, the FFR occurs at approximately 6- 10 ms after the onset of the stimulus, whereas the CM occurs simultaneously with the stimulus. Thus, the longer latency of the FFR indicates that the response is originating from a neural source beyond the cochlea. In summary, results of experiments conducted on animals and humans have confirmed that the FFR recorded on the human scalp

originates in the rostral brainstem (Du, Kong, Wang, Wu, & Li, 2011; Marsh, Brown, & Smith, 1974).

Several investigators have been interested in studying which specific regions within the auditory brainstem region contribute to the FFR. The results of these studies will be discussed in more detail in the neural generators section which follows.

### **Neural Generators of the FFR**

The neural generators of the FFR have been investigated in both animals and humans. Below is a review of these studies and their results.

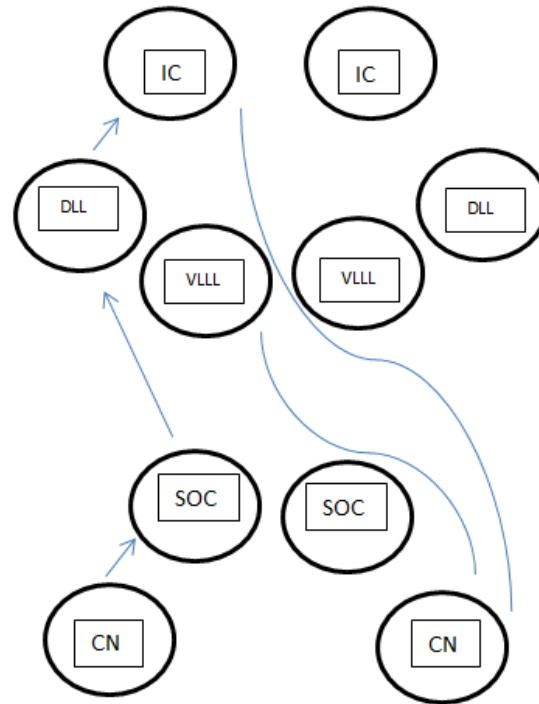
#### ***Results from Neural Generator Studies in Animals:***

In order to determine the underlying neural generators that contribute to the FFR, Marsh and colleagues analyzed FFR responses recorded at various areas in the brainstem of cats using needle electrodes placed during stereotaxic surgery (Marsh et al., 1974). These neural responses were recorded at approximately 50 different electrode placements in the auditory neural pathways of twenty seven adult cats. The anatomical locations that were of primary interest were: the Medial and Lateral Superior Olive complexes (MSO and LSO); the trapezoid body; the dorsal and ventral nucleus of the lateral lemniscus; and the inferior colliculus. The FFRs were elicited by presenting tone burst stimuli of different stimulus frequencies both monaurally and binaurally (Marsh et al., 1974). Marsh et al. (1974) reported that there were “two discrete pathways for the conduction of the frequency following responses (FFR) from the cochlear nucleus to the inferior colliculus”. The first was a contralateral pathway which began at the cochlear nucleus and terminated directly in the contralateral ventral lateral lemniscus, and contralateral inferior colliculus (Marsh et al., 1974). In contrast, the ipsilateral pathway began at

the cochlear nucleus, synapsed on the superior olivary complex, and projected towards the dorsal nucleus of the lateral lemniscus, and inferior colliculus on the same side (Marsh et al., 1974).

These ipsilateral and contralateral pathways can be seen in the Figure 6 below.

Marsh and colleagues also reported that if a lesion exists in the SOC, it does not affect the FFR recorded along the contralateral pathway, since the contralateral pathway bypasses the SOC and the neurons directly synapse on the lateral lemniscus and inferior colliculi (Marsh et al., 1974). This current evidence as well as an earlier finding that Marsh et al. (1970) demonstrated that cooling of the cochlear nucleus affects the ipsilateral pathway only and eliminated the response from the SOC on that side collectively supports the idea that there are two discrete pathways for conduction of the FFR.



*Figure 7.* Ipsilateral and contralateral pathways of the FFR neural generators. Ipsilateral pathway consists of the following structures: cochlear nucleus (CN), superior olivary complex (SOC), the Dorsal Lateral Lemniscus (DLL), and the inferior colliculus. The contralateral pathway consists of the cochlear nucleus, the Ventral Lateral Lemniscus (VLL), and the inferior colliculus (Marsh et al., 1974).

In the mid-1970s, Smith and colleagues also conducted a study which compared FFRs recorded from cats' scalps (far-field recording) to near-field recordings obtained from various brainstem nuclei in cats in order to identify the neural generators contributing to the scalp-recorded response (Smith et al., 1975). In this study, a 500 Hz tone was used to elicit the FFR in four cats. Latency measurements were compared between near-field recordings obtained at the CN, SOC, and IC to those obtained from far-field recordings on the animal's scalp. Smith et al. (1975) reported that the FFR onset latency for the far-field scalp recording was approximately 6

ms. In contrast, the FFR onset latency values obtained from the near-field recording increased in value from the cochlear nucleus to the inferior colliculus. The FFR onset latency at the inferior colliculus (5.4 ms) was the most similar to the scalp FFR latency (6 ms). Thus, Smith and colleagues (1975) concluded that the IC was the neural generator of the FFR in these animals.

In this study, Smith et al. (1975) also confirmed that the IC was the primary neural generator of the FFR in the cat by demonstrating a substantial reduction in the amplitude of the scalp-recorded FFR when the IC was cooled using a reversible blocking technique. In contrast, the FFR recorded at the other neural generator sites, such as the SOC and CN, were unaffected by the cooling technique (Smith et al., 1975). Collectively these results suggest that the IC is a neural generator of the FFR in cats.

### ***Results from Neural Generator studies in Humans***

Similar to the animal studies, Sohmer, Pratt, and Kinarti (1977) also found that the inferior colliculus was a neural generator of the FFR in humans. Sohmer and colleagues elicited the FFR using tone bursts and recorded the FFR in individuals with and without brainstem lesions. A comparison between the two populations revealed that the FFR was absent in those participants who had known “upper brainstem lesions” and present in individuals with a normal functioning brainstem. Results from this study confirmed that the upper brainstem lesion prohibited the FFR from being recorded, suggesting that the neural generator for the FFR was located at the rostral brainstem level (Sohmer et al., 1977).

Other far-field recording experiments on humans have used complex stimuli to elicit and study the neural generators of the FFR. In a study conducted by Galbraith (1994), two FFRs were recorded using two different electrode montages. One montage consisted of recordings obtained

between two inverting electrodes placed on the right and left earlobes and a non-inverting electrode placed at Fz (i.e., a horizontal electrode montage). The other electrode montage consisted of a non-inverting electrode located at the vertex of the head (Cz) and the second, inverting electrode on the ear lobe (i.e. vertical electrode montage) (Galbraith, 1994). In this study, FFRs were recorded to a 200 Hz pure tone stimulus and a complex stimulus with a missing fundamental using both electrode montages. (Note: a stimulus with a missing fundamental occurs when the frequencies present in the stimulus are all higher than the missing fundamental frequency). The intensity of both of the stimuli was 70 dB re: SPL. These investigators reported that the 200 Hz stimulus elicited a response in both the vertical and horizontal recordings. On the other hand, when the FFR was recorded to the complex stimuli, the FFR was present using the vertical electrode montage and was absent using the horizontal electrode montage. The absence of the FFR to the complex stimuli using the horizontal channel indicates that this type of complex information is not processed in the peripheral auditory system, but rather in the central auditory system. The horizontal electrode montage assesses electrophysiologic responses originating in the peripheral auditory system only. The FFR to the complex stimuli recorded using the vertical montage had a latency of 4.5 ms, which is consistent with later waves of the ABR. This finding suggests that the FFR to the complex stimuli is processed at the level of the brainstem (Galbraith et al., 2000). Based on this collective evidence, Galbraith and colleagues concluded that the FFR recorded in humans with the vertical electrode montage reflects neural generators that are at the level of auditory brainstem.

Furthermore, Chandrasekaran and Kraus (2010) stated that there are “fewer neural structures higher in the auditory pathway that are able to synchronize their firing patterns to the stimulus as compared to the auditory structures early in the pathway”. Specifically, they reported

that at the level of the Medial Geniculate Body (MGB) and above, the phase locking capabilities of auditory neurons are poor, whereas there are many “synchronized units” located at the level of the inferior colliculus. Since, the FFR is a “phase locked” response to the stimulus, it can be concluded that the neural generators for this response are located in the lower areas of the brainstem.

In summary, both animal and human FFR studies have suggested that the phase locked FFR originates in the auditory brainstem region. Collectively, the evidence from these studies conducted on both animals and humans have supported that the scalp recorded FFR is a combination of neural activity originating from the CN, SOC, LL, and IC within the auditory brainstem region (Chandrasekaran & Kraus, 2010). The following section is an overall discussion of the FFR.

### **Basics of the FFR**

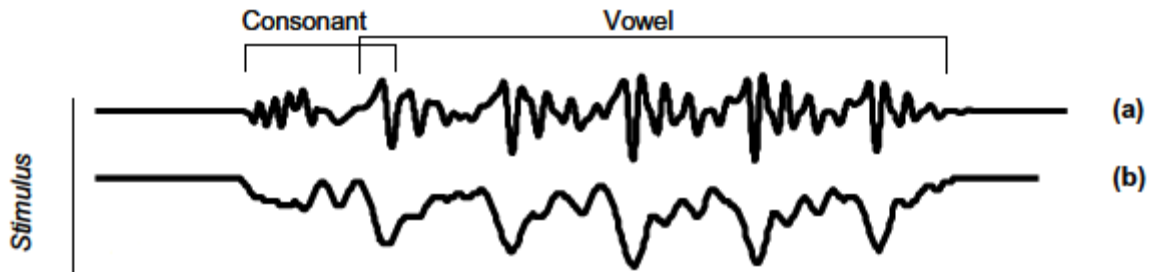
The FFR provides information about how well the brainstem is encoding pitch within an auditory stimulus. During FFR recordings, surface electrodes are placed at various locations on the scalp which measure the “activity of synchronized neural discharges” to an auditory stimulus (Aiken & Picton, 2008). FFRs are typically evoked using a complex stimulus such as speech, but can be elicited using other stimuli. Researchers have used other stimuli as well such as pure tones, iterated rippled noise, digitally created vowels, and CV syllables to elicit the FFR (Krishnan, 1999; Smith et al., 1975; Song, Banai, Russo, & Kraus, 2006; Swaminathan, Krishnan, Gandour, & Xu, 2008). The neural response from the brainstem recorded using the FFR reflects the frequency of the incoming stimulus (Hairston, Letowski, & McDowell, 2013; Kraus & Nicol, 2005), as the firing patterns of the neurons in the rostral brainstem are phase-



locked to the frequency content of the stimulus. Hence, the FFR provides an effective physiologic window into studying the phase locking capabilities of auditory neurons that are located in subcortical areas of the brain. The neurons in the brainstem can “phase lock” for stimuli up to 1500 Hz (Galbraith et al., 2000). Neural phase-locking ability decreases at higher stimulus frequencies (Galbraith et al., 2000). At frequencies beyond this limit, neural adaptation will occur which ultimately impairs the neural phase locking capabilities of the brainstem (Kvale & Schriener, 2004).

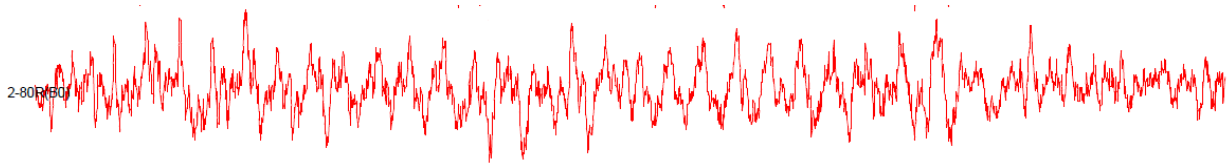
The FFR contains two pieces of information which are embedded in the temporal waveform. These two pieces of information are the envelope of the stimulus and the fine structure characteristics of the stimulus. The envelope of the stimulus is related to the fundamental frequency, whereas the spectral characteristics or fine structure of the stimulus are related to the formant frequencies (Aiken & Picton, 2008).

The initial temporal waveform obtained during FFR recordings is a summed or averaged response to both the spectral characteristics and the envelope of the stimulus (Aiken & Picton, 2008). An example of the different spectral characteristics and the envelope of the stimulus /ba/ is pictured below.

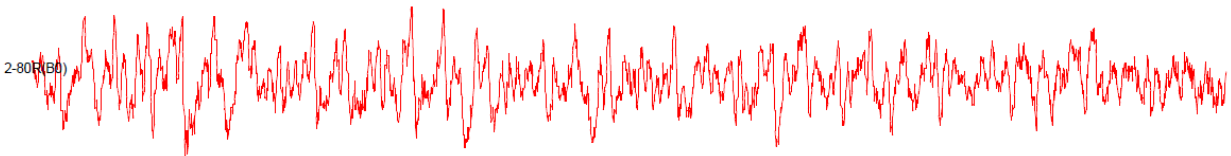


*Figure 8.* The overall spectral characteristics of the stimulus (a) and the envelope of the response (b) represented in a temporal waveform (Y axis is amplitude and X axis time) (Akhoun et al., 2008a). The consonant is the onset of the stimulus and the vowel is the sustained part of the stimulus

It is recommended that the FFR be recorded to two different stimulus polarities (i.e. rarefaction and condensation) to allow researchers to differentiate between the response to spectral properties of the stimulus and the overall envelope of the stimulus. Generally the FFR is first recorded separately to both condensation and rarefaction phases of the stimulus. It can also be recorded to the alternating phase of the stimulus and then subsequently separated into the averaged rarefaction and condensation waveforms. These averaged waveforms which contain the response to a single polarity stimulus are then added or subtracted together to distinguish between the response to the envelope and the response to the fine structure of the stimulus (Skoe & Kraus, 2010). The individual temporal waveforms recorded to the condensation and rarefaction phases of the stimulus are pictured below.

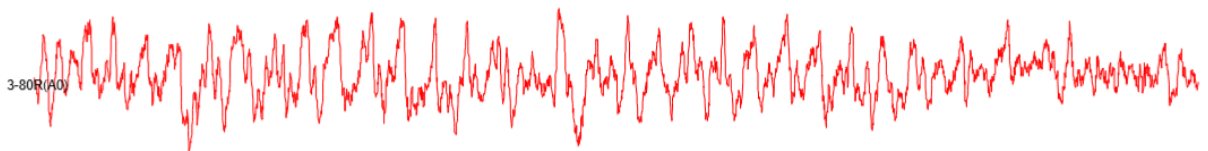


*Figure 9.* Waveform recorded to the condensation polarity of the stimulus



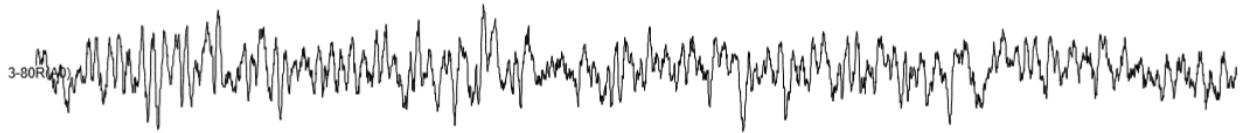
*Figure 10.* Waveform recorded to the rarefaction polarity of the stimulus

The addition of the rarefaction and condensation waveforms yields a summed waveform that represents the envelope of the stimulus (Skoe & Kraus, 2010). Below is an image of the FFR in the temporal domain after addition of the two separately recorded waveforms to the condensation and rarefaction stimulus.



*Figure 11.* Condensation + Rarefaction waveform

On the other hand, a subtraction of the two waveforms recorded using a condensation and rarefaction polarity yields a waveform that solely follows the spectral components or TFS of the speech stimulus (Aiken & Picton, 2008). Below is an image of the FFR in the temporal domain after the two waveforms were subtracted.

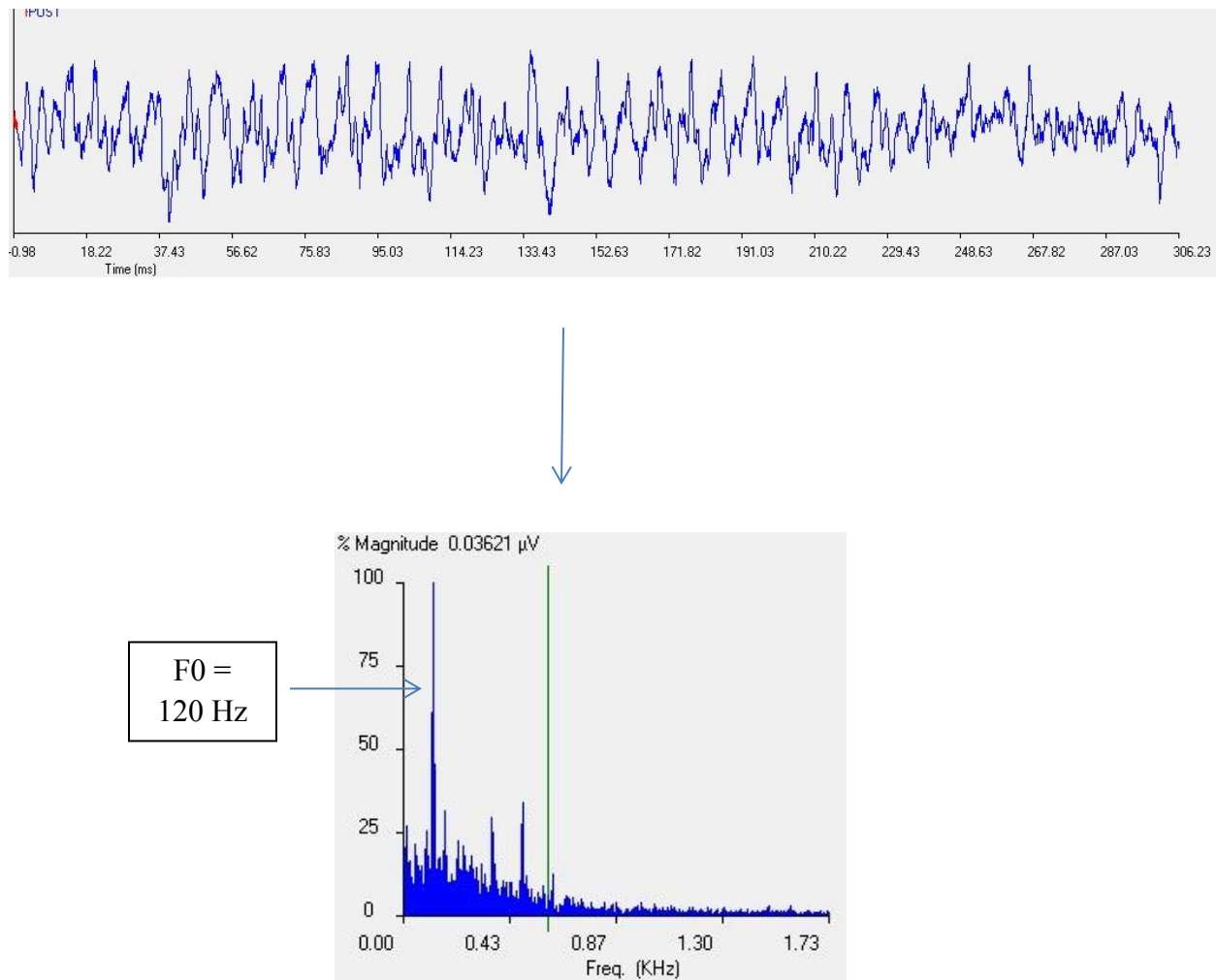


*Figure 12.* Condensation – Rarefaction waveform

Once the FFR is broken down into the two components (spectral/fine structure FFR and envelope FFR) in the temporal domain, these two responses can be analyzed in the spectral domain using the Fast Fourier Transform (FFT) to determine where the energy is present in the FFR spectrum. The FFT allows the researcher to view the FFR in the frequency domain and determine if the response contains energy at the fundamental frequency and at the formants of the stimulus (Russo, Nicol, Musacchia, & Kraus, 2004). As the FFR reflects neural phase locking to both envelope and TFS cues, the spectral analysis of the response to the speech stimulus should indicate robust energy at frequencies corresponding to the fundamental frequency and at the formant frequencies of the stimulus. A direct comparison between the acoustic properties of stimulus ( $F_0$ ,  $F_1$ , and  $F_2$ ) and the FFT of the response will reveal how well the brainstem encoded the speech stimuli.

Below is an example of an FFT analysis applied to the summed envelope FFR. The FFR was elicited to a speech stimulus /u/, which has a fundamental frequency of 120 Hz and a formant frequency ( $F_1$ ) of 360 Hz. The temporal waveform is the summed waveform resulting from the addition of the waveforms to the rarefaction stimulus and the condensation stimulus. An FFT is then applied to this summed waveform which displays a prominent band of energy present at the fundamental frequency ( $F_0$ ) of 120 Hz. This band of energy represents the envelope of the stimulus. Since the largest band of energy is present at the fundamental

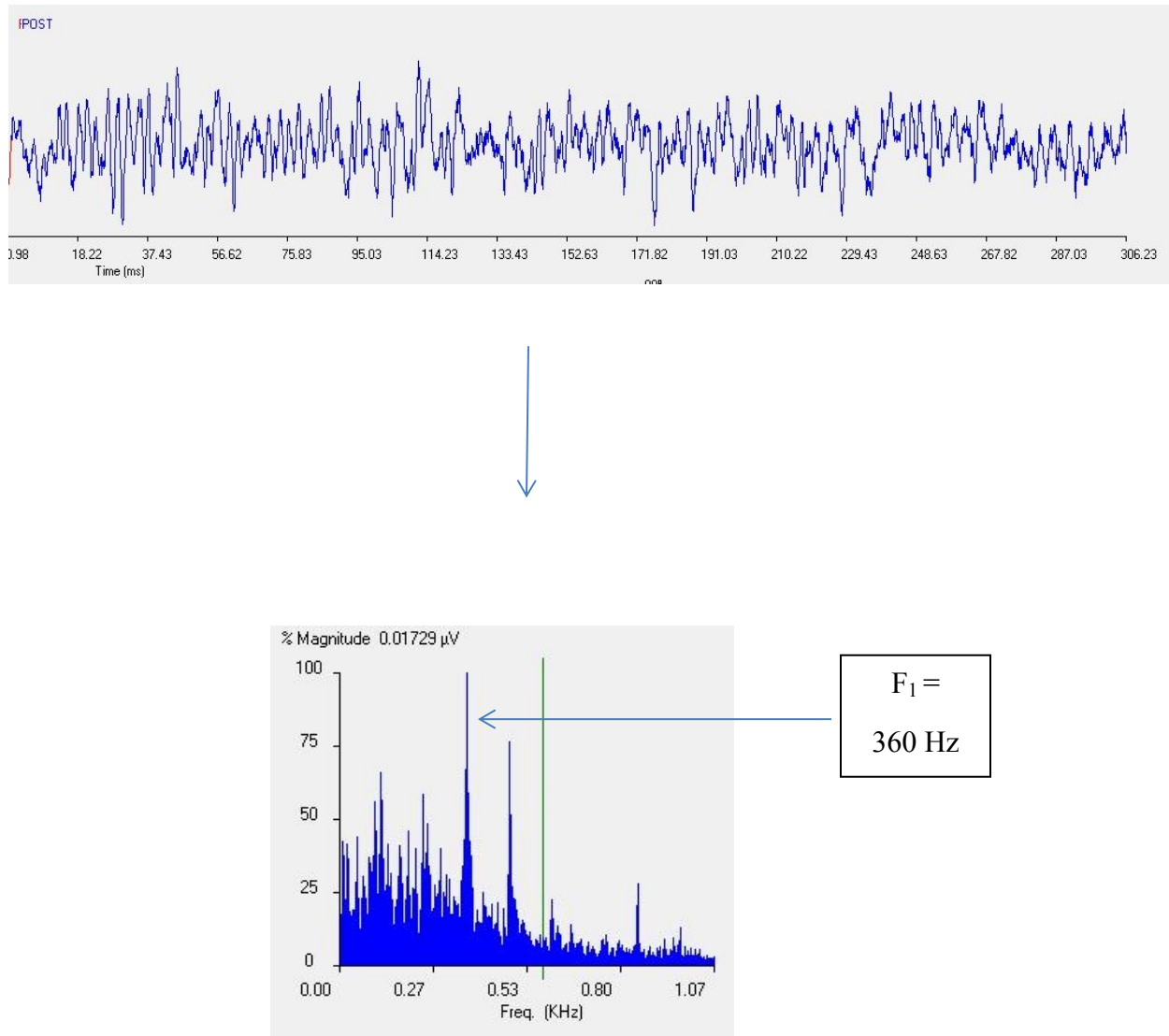
frequency of the stimulus, this indicates that the neurons in the sub cortex have encoded the envelope of the stimulus



*Figure 13.* Summed envelope FFR and FFT

The FFT was also applied to the difference waveforms representing the TFS of the stimulus. Again, the FFR was elicited using a /u/ stimulus with the same fundamental and first formant frequencies, 120 Hz and 360 Hz, respectively. The temporal waveform is the subtracted waveform resulting from the subtraction of the waveforms to the rarefaction stimulus and the

condensation stimulus. An FFT is then applied to this subtracted waveform which displays a prominent band of energy at the first formant frequency ( $F_1$ ). The greater energy present at  $F_1$  represents the fine structure of the stimulus. Therefore, the fine structure of the stimulus is efficiently coded at the level of the auditory subcortex.



*Figure 14.* Difference waveforms representing the TFS of the stimulus and FFT

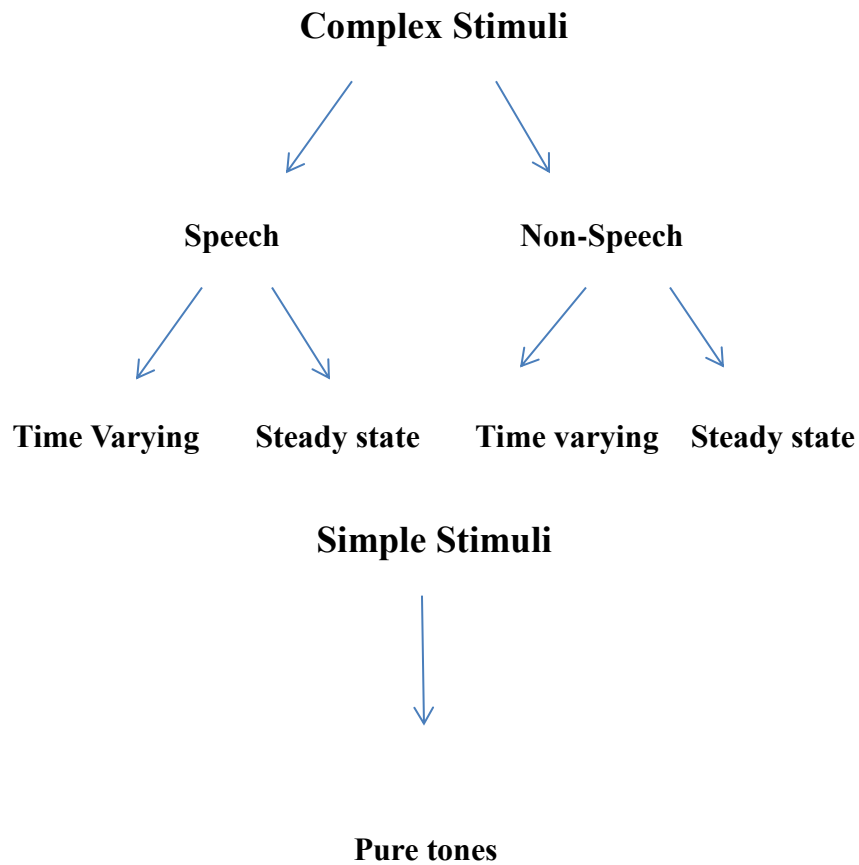
Technical parameters can alter the information obtained during Auditory Evoked Potential recordings. Several important technical parameters include stimulus, recording, and

subject parameters. Stimulus parameters describe the characteristics of the stimulus such as the type, intensity, rate, and polarity of the stimulus. For the purpose of this study all three technical parameters will be discussed in further detail.

## **Stimulus Parameters**

### ***Type of Stimulus***

The brainstem FFR can be elicited in response to both simple as well as complex stimuli. A simple stimulus consists of one sinusoidal waveform or pure tone. It was the chosen stimulus in many of the early FFR studies (Marsh et al., 1974; Moushegian et al., 1973; Smith et al., 1975). On the other hand, a complex stimulus is composed of several different waveforms or frequencies (Emanuel, 2009). Moushegian and colleagues found that pure tones ranging from 250 Hz to 2000 Hz elicited neural potentials which were phase locked to the stimulus or the auditory neurons fired at whole number multiples of the stimulus as shown by the FFR. More recently, most of the FFR studies conducted in the last couple decades have collected FFR data using complex stimuli. A complex stimulus can be a speech stimulus or a non-speech stimulus. Speech and non-speech stimuli are further characterized as “time varying” stimuli or “steady state” stimuli re: their fundamental frequency. Recall from the “basic acoustics” section that the fundamental frequency of the stimulus is the rate at which the vocal chords vibrate. Speech and non-speech stimuli can be “time varying” indicating that the fundamental frequency or formants changes over the length of the stimulus; or “steady state” where the fundamental frequency and formants remains stable over the entire duration of the stimulus (Swaminathan, Krishnan, Gandour, & Xu, 2008). Below is a diagram of the different types of complex and simple stimuli.

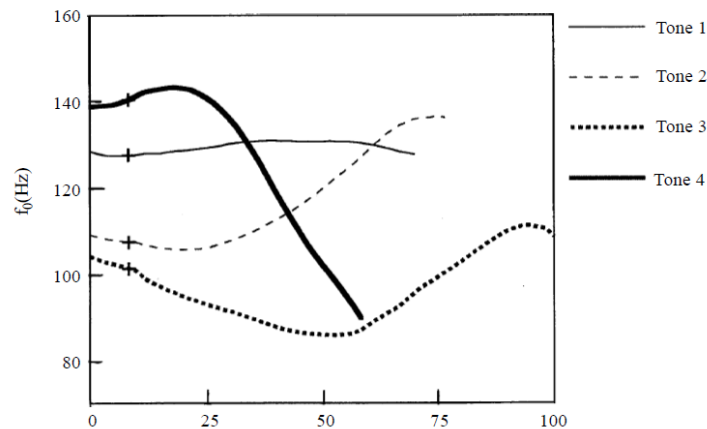


*Figure 15.*Types of complex and simple stimuli

An understanding of different types of stimuli is important because the FFR “preserve[s] the frequency characteristics” of both simple and complex stimuli (Swaminathan et al., 2008). An example of a complex non speech stimulus used in the FFR literature is iterated rippled noise. Iterated rippled noise (IRN) is broadband noise that can be manipulated to create a perception of pitch. In a study conducted by Swaminathan and colleagues (2008), a steady state and time varying IRN was used to elicit the FFR. The steady state IRN had a constant fundamental frequency of 240 Hz. On the other hand, the fundamental frequency changed from 101Hz to 132 Hz for the time varying IRN. Swaminathan and colleagues concluded that both



steady state and time varying IRN stimuli can “elicit an FFR with fine-grained measures of pitch perception”. On the other hand, an example of a speech stimulus that has been used to elicit the FFR that was both time varying and steady state is the Mandarin Chinese syllable. In a study conducted by Krishnan, Xu, Gandour, and Cariani (2005), four different lexical tones of the syllable /yi/ were used to evoke the FFR. The pitch of the lexical tone /yi/ varied over time. The first tone exhibited a stable F0 throughout the duration of the stimulus (a steady state stimuli), while the remaining lexical tones were all time-varying in nature. Specifically, the second and third lexical tones had rising and falling pitch contours respectively. The fourth lexical tone had a falling pitch contour. Figure 2 below is a graphical representation of the F0 changing over time for all four lexical tones. In conclusion this study shows that neurons within the brainstem can follow both steady-state and time-varying frequencies.



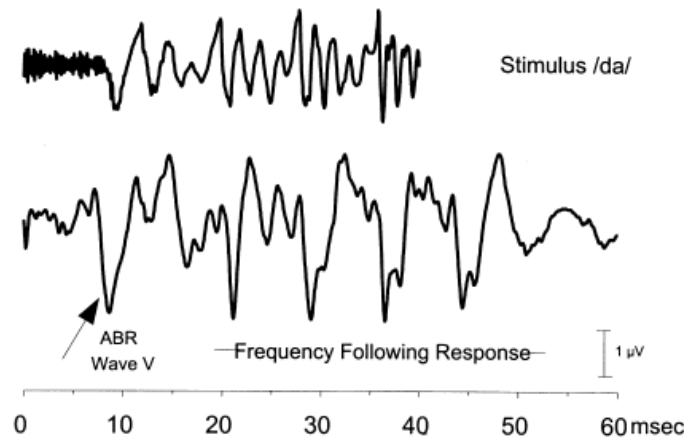
*Figure 16.* F0 changing over time for four lexical tones

Another example of a non-speech stimulus is a “two tone approximation” that simulates the frequencies of a complex vowel (Krishnan, 1999). For example, in the study conducted by

Krishnan (1999) two separate stimulus frequencies were combined with a formant frequency structure similar to the vowel /ɔ/. The rationale for using a “two tone approximation” of a vowel was to determine if the FFR could be elicited to a non-speech stimulus that resembled a speech stimulus in spectral structure.

The remainder of this section will focus on the FFR studies that have used speech stimuli, since a speech stimulus will be used to elicit the FFR for this current study. Speech can be naturally produced or digitally created (Aiken & Picton, 2008; Cunningham, Nicol, Zecker, Bradlow, & Kraus, 2001). In 1979, Young and Sachs summarized the results from a study that used a synthesized steady state speech stimulus to elicit single unit auditory nerve responses in cats using a near field recording. The three different steady state vowels used were /i/, /a/, and /e/, all with a fundamental frequency of 128 Hz. Young and Sachs concluded that there were “synchronized responses” to the formant frequencies and the fundamental frequency of the stimuli. Thus, indicating that speech stimuli can be used to demonstrate the phase locking capabilities of the auditory system. Naturally produced vowels like /i/ and /a/ have been successfully used to elicit the FFR (e.g., Aiken & Picton, 2008). Aiken and colleagues recorded the FFR in seven women and three men using these two speech stimuli. Aiken and Picton (2008) concluded that FFR responses were phase locked to the stimulus and preserved the fundamental frequencies and harmonics of the stimulus. This finding provides additional support that speech stimuli can be successfully used to elicit the FFR, the phase locked response at the brainstem level. Additional speech syllables which have been used to elicit a robust FFR include /ada/ and /aga/ as well as the Mandarin syllables mentioned above (Cunningham et al., 2001; Skoe & Kraus, 2010).

A common speech stimulus used to elicit the FFR is the Consonant Vowel (CV) stimulus such as /da/ (Song et al., 2006). CV stimuli elicit both a transient and a sustained response. The transient and sustained response is characterized based on the timing of the response in relation to the stimulus (Skoe & Kraus, 2010). The transient portion of the FFR occurs at the onset or offset of the stimulus and is present for only a tenth of a millisecond (Song et al., 2006). In contrast, the sustained portion of the FFR occurs throughout the length of the stimulus and is phase locked or “time locked” to the period of the stimulus (Skoe & Kraus, 2010). The sustained response is a response to the “continuous acoustic features” of the stimulus, such as the formants, fundamental frequency, and harmonic frequencies (Skoe & Kraus, 2010). An example of a consonant vowel stimulus is /da/. The consonant /d/ is the rapid, transient portion of the stimulus eliciting a quick response at its onset. In contrast, the vowel /a/ is a “sustained period signal” that elicits the sustained, synchronous brainstem response (Song et al., 2006, p. 239). As seen in the figure below a response occurs at the onset of the stimulus /da/ and the response continues to follow the period of the stimulus. The FFR captures both the fundamental and formant frequencies of the speech stimulus /da/. Words and phrases have also been used as speech stimuli for the FFR, but are not used as routinely as CV syllables (Skoe & Kraus, 2010).



*Figure 17.* The FFR mimics the period the stimulus /da/. The figure was extracted from Cunningham et al. (2001)

As the major focus of the current study is on subcortical encoding of speech in humans in adverse listening conditions, and given the success of speech stimuli in eliciting robust FFRs, the stimulus that will be used in the current experiment is the /u/ stimulus.

### ***Intensity of Stimulus***

Several studies have examined the effects of stimulus intensity on subcortical encoding of auditory stimuli as represented by the FFR. Davis and Hirsh (1976) recorded the FFR to a 500 Hz tone burst in 21 adult individuals mostly with normal hearing. Davis and colleagues found that the threshold for the FFR is approximately 40 dB SPL. They reported that at this stimulus intensity, the temporal waveform of the FFR was not well defined and the latencies of the waves were prolonged. Davis and Hirsh (1976) concluded that the FFR “is rarely clear below 50 dB SPL”. These results established that the FFR must be recorded at at least 50 dB SPL to obtain a clearly defined response.

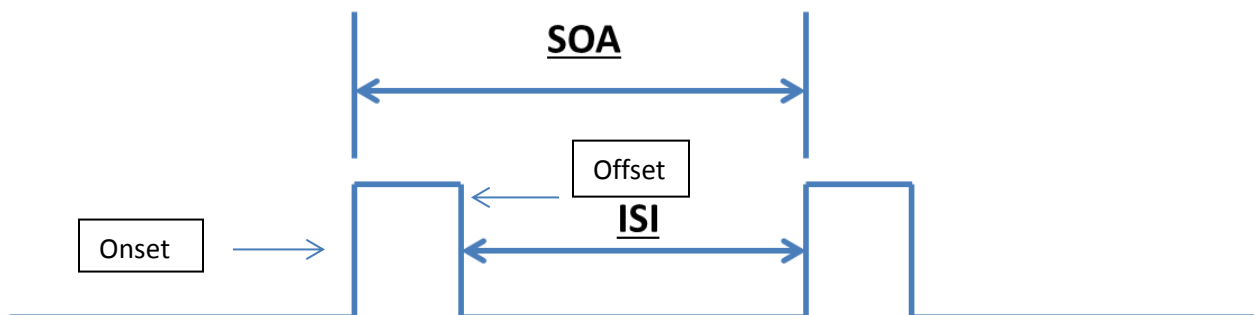
Krishnan (2002) found similar results when they elicited the FFR on normal hearing adults with different vowels (/ʊ/, /ɑ/, /ɔ/ ) at different intensity levels of 55, 65, 75, and 85 dBnHL. Krishnan (2002) found an “overall reduction in response amplitude of the FFR as the stimulus intensity decreased”. The FFR exhibited “stronger” or more robust phase locking capabilities to the formant frequencies of the stimulus at louder intensities (Krishnan, 2002). Overall, Krishnan and colleagues concluded that the brainstem encodes the formant frequencies of the stimulus best at intensity levels greater than 65 dB nHL.

Similarly, Akhoun et al. (2008) elicited the FFR in 23 adult females and males with normal hearing thresholds using a /ba/ stimulus. This stimulus was presented at various stimulus intensities ranging from 0 to 60 dB SL. Akhoun and colleagues reported that as stimulus intensity increased, the latency of the FFR decreased by approximately 1.4 ms for every 10 dB increase (Akhoun et al., 2008). These investigators reported that in general, a decrease in the intensity of the stimulus increased the latency of the FFR (Akhoun et al., 2008).

Collectively, the results of these three studies demonstrate that a high intensity stimulus is necessary to elicit a robust FFR. Given this evidence, the speech stimuli used to elicit the FFR are usually presented at suprathreshold levels ranging from 60 to 85 dB SPL (Skoe & Kraus, 2010). The FFR has been found to be best defined at suprathreshold levels around 60 dB SL and is essentially flat at intensities near threshold (Skoe & Kraus, 2010). The intensity of the stimulus also affects the latency of the FFR as seen in the research conducted by Akhoun and colleagues (Skoe & Kraus, 2010). Therefore, in the present study, the FFR will be recorded at a suprathreshold level of 75 dB in order to maximize the identification and amplitude of the response.

### ***Rate of Stimulus***

Audiologists should be aware of the presentation rate of the stimulus because rate can alter the FFR (Krizman, Skoe, and Kraus, 2010). Stimulus rate is the number of stimuli presented per second. The presentation rate of the stimulus can change the brainstem's ability to encode certain frequencies of the stimulus, which is reflected in the FFR. The rate of the stimulus presentation can be determined by the inter-stimulus interval (ISI) or stimulus onset asynchrony (SOA) (Skoe & Kraus, 2010). ISI is defined as the silent period, measured in milliseconds, between the offset of a stimulus and the onset of the next stimulus (Skoe & Kraus, 2010). ISI is the reciprocal of stimulus rate, for example an ISI of 100 ms is equivalent to a stimulus rate of 10 stimuli per second ( $1000 \text{ ms} / 100 \text{ ms} = 10 \text{ stimuli per second}$ ). Skoe and Kraus (2010) stated that a review of the FFR literature revealed that the ISI should be at least 30 % of the total duration of the stimulus. The rate of the stimulus presentation can also be defined as the stimulus onset asynchrony (SOA). SOA is the duration between the “onset of one stimulus and onset of the next stimulus. As can be seen below the SOA occurs from the onset of the first stimulus to the onset of the following stimulus, whereas the ISI occurs between the offset of the first stimulus and the onset of the second stimulus.



*Figure 18.* Comparison between SOA and ISI

Several factors should be considered when selecting an appropriate ISI or stimulus rate. One factor that needs to be considered when selecting an appropriate stimulus rate is that the FFR response should return to baseline in between stimulus presentations. Neural adaptation is a decrease in neural response that occurs when the brain is unable to recover between successive stimuli (Kvale & Schriener, 2004; Skoe & Kraus, 2010). Therefore, the stimulus presentation rate should be chosen so that the neurons contributing to the FFR are able to fire at their maximum potential each time the stimulus is presented. A second important factor is that the stimulus rate can affect how well the FFR can phase lock to high frequency components of a stimulus (Krizman, Skoe, & Kraus, 2010). Krizman et al. (2010) conducted a study to evaluate the effects of three stimulus presentation rates (slow: 6.9 Hz, standard: 10.9 Hz, and fast: 15.4 Hz) on the FFR. The ISI for the slow, standard, and fast rate speech stimuli would be calculated as approximately 144 ms, 91 ms, and 64.9 ms, respectively. The stimulus used in the study was a five formant synthesized speech sound. Spectral analysis of this stimulus revealed the energy at the fundamental frequency ranged from 103 to 125 Hz, while the energy at the formant frequencies was as follows:  $F_1$  ranged from 220 to 720 Hz,  $F_2$  from 1700 to 1240 Hz,  $F_3$  from 2580 to 2500 Hz, and  $F_4$  and  $F_5$  are constant at 3600 and 4500 Hz, respectively. Krizman et al. (2010) reported that the FFR to the fundamental frequency and low frequency formants (180 Hz to 410 Hz) of the stimulus were unaffected by stimulus rate. However, the high frequency formants ranging from 411 Hz to 755 Hz were not encoded as efficiently as stimulus rates increased from 6.9 Hz to 15.4 Hz. Furthermore, the energy in the neural response at the harmonics ranging from 756 Hz to 1130 Hz was larger as the stimulus rate slowed (Krizman et al., 2010). Overall, the brainstem's ability to follow the low frequency energy present in the stimulus was unaffected with changes in stimulus rate, whereas the brainstem's ability to follow

the high frequency energy present in the stimulus worsened with increasing stimulus rate. Clearly, this experiment indicates that stimulus rate should be carefully selected to obtain an optimal recording of the FFR (Krizman et al., 2010). Given this evidence a stimulus rate of 3.13 /second will be used for this study.

### ***Polarity of Stimulus***

Another important stimulus parameter is the polarity of the stimulus. Polarity of the stimulus refers to the “positive or negative deflection of the sound wave relative to the baseline”. The two types of stimulus polarities used to record the FFR are condensation and rarefaction. Condensation stimuli result in a positive deflection of the sound wave relative to the baseline; whereas rarefaction stimuli result in a negative deflection from the base line (Skoe & Kraus, 2010). The diaphragm of the microphone producing a condensation stimulus polarity will move inward. Physiologically this results in an inward movement of the stapes and upward deflection of the basilar membrane (Moller, 2006). Conversely, the diaphragm of the microphone producing a rarefaction stimulus polarity will move outward, resulting in the outward movement of the stapes and a downward deflection of the basilar membrane (Moller, 2006). Stimulus polarity is important to understand when recording the FFR because this response is separately recorded to each of these stimulus polarities in order to “tease apart” the two primary components of the FFR ( i.e., the envelope and the fine structure of the response) (Skoe & Kraus, 2010). Specific details regarding how a change in stimulus polarity is used to tease apart the individual components of the FFR was discussed in detail in “What is the FFR” section of this literature review. For this study, the FFR will be recorded using both stimulus polarities and FFTs will be applied to both averaged waveforms to determine the fundamental and formant frequencies present in the FFR.



## **Recording Parameters**

Recording parameters in AEP literature are a set of guidelines that are used to regulate collection of the neural responses. Recording parameters define aspects related to data collection such as the analog band-pass filter settings, the length of the analysis window, the number of sweeps contributing to the averaged response, electrode montage, sampling rate, and artifact rejection. Subject parameters are the participant/patient-related aspects of the AEP recording. Two subject parameters to consider during FFR recordings are subject state and age. Both the recording and subject parameters will be discussed in the below section.

### ***Band-Pass Filter***

During the FFR, the recording electrodes measure the desired neural response (i.e. FFR) along with constant electrical activity of the brain (EEG). The amplitude of the FFR is relatively small in comparison to the background EEG being recorded. There are several different types of filters such as a high-pass filter, a low-pass filter, a band reject filter, and the band-pass filter. A high-pass filter consists of one high frequency cut off value where frequencies below the chosen value are rejected. On the other hand, a low-pass filter consists of a one low frequency cut off value where frequencies above the chosen value are rejected. A band reject filter is narrow band of frequencies between a low frequency and a high frequency value that are filtered out. An analog band-pass filter can be used to help isolate the desired response from the unwanted background noise, thereby increasing the SNR (Skoe & Kraus, 2010). A band-pass filter is a filter setting with two designated cut off frequencies. For example a band-pass filter ranging from 100 to 3000 Hz has a low frequency cut off value of 100 Hz and a high frequency cut off value of 3000 Hz (Atcherson & Stoody, 2012). The aim or goal of the band-pass filter is to

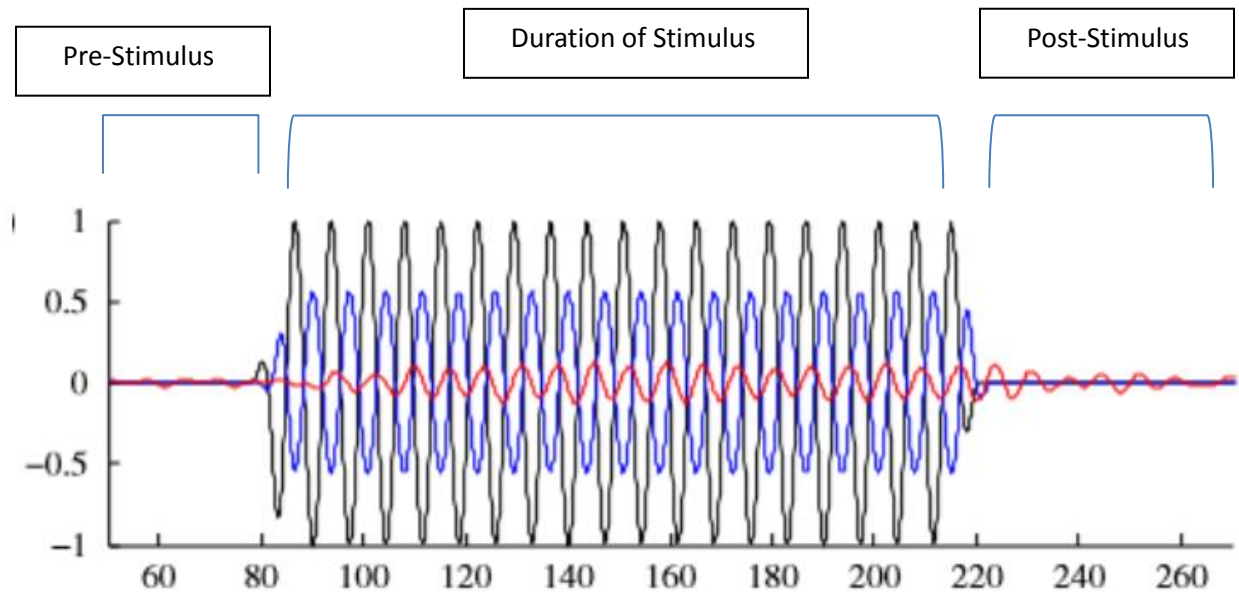
include the frequencies or energy in the desired neural response and “filter” out the noise. In other words the frequencies that are below the low frequency cut off value or above the high frequency cut off value are not included in the overall averaged response (this is in contrast to the band reject filter). Evoked potential literature has reported the use of band-pass filters with different low-pass and high-pass filter cut off values depending upon the auditory evoked potential being recorded.

The cut off values of the analog filter can be manipulated depending on where the energy is present in the response. For instance, Lee, Skoe, Kraus, and Ashley (2009) used a band-pass filter ranging from 20 – 2000 Hz to record the FFR elicited by stimuli with frequencies ranging from 93 Hz to 166 Hz. As the FFR has energy present at the fundamental frequencies and harmonics of the stimulus a band-pass filter setting from 20 – 2000 Hz was appropriate since the response energy fell within this range. Similarly, Sinha and Basavaraj (2010) used an analog band-pass filter setting of 30-3000 Hz while recording the FFR to the speech stimulus /da/ which had a fundamental frequency ranging from 105 to 121 Hz. Again these filter settings encompassed the energy present in the FFR at the fundamental and formant frequencies. Skoe and Kraus (2010) reported that band-pass filters ranging from 100 – 3000 Hz and 30- 3000 Hz are the desired band-pass filter settings for recording the FFR. According to Skoe and Kraus, the 100 – 3000 Hz band-pass filter helps capture the “high frequency transient peaks” of the response. The general consensus of these studies was that the band-pass filter settings must be set at a level such that they capture the energy present in the response. In the case of the FFR, which mimics the stimulus, these levels were determined by the F0 of the stimulus and that phase-locking which decreases with increasing in frequency. The F0 of the stimulus used in the current

study is 120 Hz and therefore an analog band-pass filter setting of 100 to 3000 Hz will be used for this study.

### ***Length of the Analysis Window***

The length of the analysis window refers to how long the FFR is being recorded. Skoe and Kraus (2010) recommended that recordings should begin prior to the presentation of the stimulus (i.e. the pre-stimulus period), through the duration of the stimulus, and after the stimulus stops (i.e. the post-stimulus period). The pre-stimulus period allows for the researcher to obtain information about the individual's resting electrical brain activity (EEG) and to establish a baseline (Skoe & Kraus, 2010). The post-stimulus period will account for the “stimulus transmission delay and neural conduction time”. The post-stimulus analysis will allow the recording to continue until the response has returned to the individual's resting state (Skoe et al., 2010). Russo and colleagues (2004) used a 70 ms analysis window when eliciting the FFR to a 40 ms long /da/ stimulus, where the pre-stimulus window was 10 ms, the duration of the stimulus related response was 40 ms, and the post-stimulus window was 20 ms long. In another example, Gong, Xu, and Sun (2013) the FFR was elicited using a 144 ms long pure tone. The recording window was set with a pre-stimulus recording length of 77 ms; however the length of the analysis window post stimulus was not noted. The figure below depicts the stimulus (black line) and the FFR (red line) throughout the duration of the recording. Note the pre-stimulus portion of the analysis window occurs from 0 to 80 ms (black and red line are flat). The analysis window extends throughout the duration of the stimulus (approximately 80 to 220 ms) and then continues for a portion of time after the offset of the stimulus (approximately 220 to 260 ms). The x-axis represents the time in (ms) and the y axis represents the amplitude of the stimulus and response in ( $\mu$ V).



*Figure 19.* Image extracted from Gong et al., (2013) showing the pre-stimulus part of the analysis window, the duration of the stimulus, and the post-stimulus part of the analysis window

Skoe and Kraus recommended that a 40 ms pre-stimulus period and a 10 ms to 50 ms post stimulus period has been recommended. In the present study, we will use a speech stimulus /u/ which has a total duration of 260 ms, with a post stimulus period of ~40 ms. In order to eliminate any onset responses from the data analysis, the first 20 ms of the response will be discarded during analysis, hence there is no need for a pre-stimulus period. Therefore, the total length of the analysis window will be 300 ms.

### ***Number of Sweeps Contributing to Averaged Response***

Sweeps refers to the number of stimuli presented during a recording session. Signal averaging is a method used during recordings of all electrically evoked potentials and enables the researcher to more easily separate the desired neural response from the background noise (Gong et al., 2013). Typically a large number of sweeps is collected in order to facilitate signal

averaging, which in turn improves the SNR of the response. The theory behind signal averaging is that background noise, such as the continuous EEG, is random whereas the desired neural response is “time locked”. When a response is time locked, it means that the response occurs at a specific latency following the onset of the auditory stimulus (Gelfand, 2001). The averaging technique in theory enhances the time-locked neural response and reduces the random background noise. For example, suppose that the stimulus is set to 2000 sweeps, which means that it repeats 2000 times during the recording session. The number of times the stimulus repeats throughout the recording sessions is also known as the number of “sweeps”. The computer will analyze and sample the electrical activity during every sweep, and average the electrical activity over 2000 sweeps. The voltage of the neural response, occurring at a specific latency after the onset of the stimulus, was averaged over 2000 sweeps, enhancing the neural response (Robinson, 2008). On the other hand, averaging the random background EEG will minimize the background noise over time, thus improving the SNR. Skoe and Kraus (2010) reported that the typical number of sweeps used when recording the FFR is 1000 to 2000 sweeps. The higher the number of sweeps the better the SNR will be (Gelfand, 2009).

### ***Electrode Montage***

The electrode montage or placement of the electrodes on the subject’s head is also an important factor determining the nature of the response obtained. The electrodes can be placed on the head in either a vertical or a horizontal montage. The vertical montage consists of a non-inverting electrode placed at Cz (the vertex) and an inverting electrode on the earlobes/mastoid. In contrast, the horizontal montage consists of electrodes placed on the right and left earlobe, where one earlobe serves as a reference (inverting electrode) and a non-inverting electrode is placed at Fz (Galbraith, 1994). Galbraith (1994) investigated the difference in FFR recordings

using a vertical electrode montage versus a horizontal electrode montage. The FFR was recorded on 10 adult subjects and was elicited using a 200 Hz stimulus and a stimulus with 200 Hz as the missing fundamental frequency. Galbraith found that a 200 Hz stimulus elicited a response in both the horizontal and vertical channels. However, the FFR in the vertical channel was delayed in comparison to the response recorded using the horizontal channel. The longer latency indicates that the vertical electrode montage is recording electrical activity later in the auditory pathway than the horizontal electrode montage. Additionally a clear defined response, to the stimulus with the missing fundamental, was only present using the vertical channel. This indicates that a complex stimulus, such as a stimulus with a missing fundamental frequency, is not encoded at this region of the auditory system where the horizontal channel is recording. In conclusion, the horizontal electrode montage records neural activity originating in the peripheral auditory system, such as the auditory nerve (Galbraith, 1994). On the other hand, the vertical electrode montage records neural activity from the brainstem, such as the inferior colliculus and lateral lemniscus (Galbraith, 1994). The FFR is most robust with the vertical montage because the FFR neural generators are located in the brainstem.

One vertical electrode montage consists of electrodes placed at Cz, the ipsilateral earlobe, and the forehead (ground). Cz acts as the active or non-inverting electrode, while the electrode placed on the earlobe is the reference electrode or inverting electrode (Skoe & Kraus, 2010). The reference electrode can be placed either on the ear or the mastoid. Krishnan, Bidelman, Smalt, Ananthakrishnan, and Gandour (2012) used an electrode montage consisting of electrodes placed at Fz, M1, M2, C7, and Fpz. The electrode montage that will be used for the present study is as follows: Fz (non-inverting), Fpz, M1 and M2 (inverting), and C7 (inverting). Fpz will act as the ground for this experiment. This montage will be used for this study because it will effectively

record the neural activity at the level of the brainstem. The figure below depicts the International 10-20 system of electrode placement.

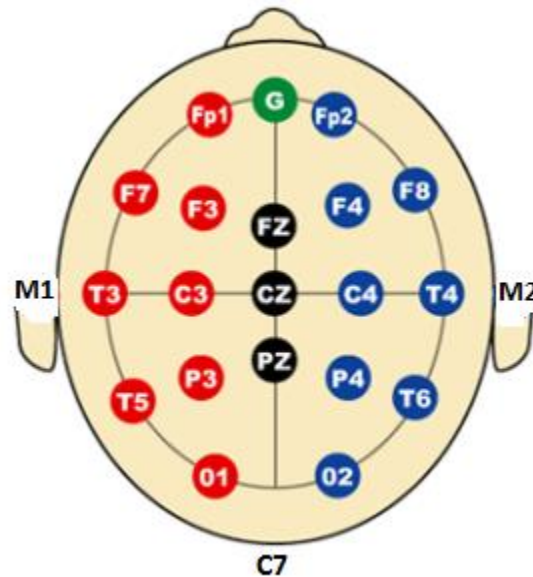
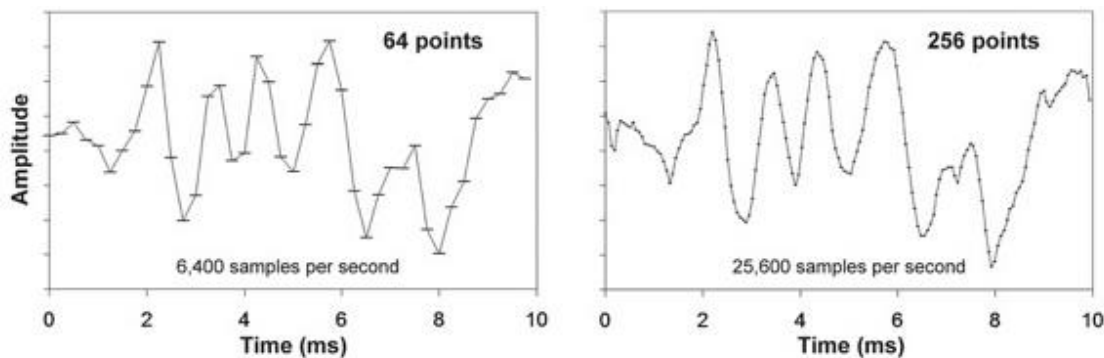


Figure 20. Electrode Placement (International 10-20 system)

### ***Sampling Rate***

Sampling rate is the number of times the neural response is sampled over the length of the analysis window (Atcherson & Stoody, 2012). The sampling rate is calculated by dividing the number of sample points by the length of the analysis window. For example, if the analysis window is 20 ms (.02 s) long and the neural activity is sampled at 1000 different points, then the sampling rate per second is 50,000. The figure below depicts a neural response over time with two different sampling rates of 6,400 per second and 25,600 samples per second.



*Figure 21.* The above figure visually depicts the neural response over a period of 10 ms. The left figure depicts 64 different sample points, whereas the right figure depicts 256 different sample points (the dots along the solid line indicate the sampling point) The sampling rate per second is 6,400 and 25,600 respectively ( $64/.01 = 6400$  and  $256/.01 = 25,600$ ). (Figure extracted from Atcherson & Stoodly, 2012)

Sampling rates are determined by stimulus frequencies (Skoe & Kraus, 2010). As a general rule the sampling rate should be twice the highest frequency present in the stimulus, which is also known as the Nyquist frequency (Skoe & Kraus, 2010). The Nyquist theory ensures that the neural activity is sampled enough to “increase temporal precision of the recording and allow for finer differentiation of response peaks” (Skoe & Kraus, 2010). For example, in a study conducted by Banai et al. (2009) the sampling rate was 6,857 Hz. The stimulus used consisted of a fundamental frequency ranging from 103 to 125 Hz. The highest frequency of the stimulus was approximately 3600 Hz (fourth formant). Therefore, the sampling rate is approximately twice the highest frequency of the stimulus or close to the Nyquist frequency. Skoe and Kraus (2010) stated that the FFR is usually recorded using sampling rates ranging from 7,000 Hz to 50, 000 Hz. Sampling rates well above the Nyquist frequency are also



commonly used during FFR recordings (Skoe & Kraus, 2010). This oversampling is recommended for a precise representation of the response. The following table lists the sampling rates used in several studies. As seen below these sampling rates vary from approximately 7,000 to 45, 000 Hz. For the current study, a sampling rate of greater than 6,000 Hz will be used to record the FFR.

*Table 1*

*Sampling Rates for FFR Studies*

Sampling Rates for FFR Studies
6,857 Hz (Banai et al., 2009)
10,000 Hz ( Smalt et al., 2012)
25,000 Hz (Krishnan, 2002)
44,000 Hz (Clinard et al., 2009)
20,000 (Russo et al., 2004)

***Artifact rejection***

During the FFR recording artifacts such as stimulus artifact, electrical artifact, myogenic artifact, and the cochlear microphonic can interfere with the neural response (Aiken & Picton, 2008). Stimulus artifact occurs when the recording electrodes record the electrical signal that is delivered to the transducer which is then converted into an acoustical signal (Hall, 2007). The stimulus artifact typically occurs milliseconds after the onset of the stimulus and is identical to the “stimulating waveform” (Skoe & Kraus, 2010). Since the stimulus artifact mimics the response, it should be appropriately controlled. Two methods can be used to reduce the

probability of the stimulus artifact interfering with the neural response. Stimulus artifact can be reduced by ensuring that the headphone and electrode wires are not overlapping (Hall, 2007). Another method used to eliminate stimulus artifact is using “electromagnetically shielded headphones” (Skoe & Kraus, 2010).

Electrical devices in the room where the neural response is being recorded can emit electrical noise that may contaminate the desired response (Hall, 2007). The frequency of the electrical noise is centered at 60 Hz and can appear as a sinusoidal waveform contaminating the neural response. One method to reduce electrical artifact is remove unnecessary electrical equipment from the room where the recordings are being conducted (Skoe & Kraus, 2010). Another method used to help eliminate the 60 Hz electrical noise is to ensure that the stimulus rate is odd and does not divide into 60 Hz evenly (Hall, 2007).

The cochlear microphonic, just like the FFR, mimics the stimulus waveform and therefore could potentially interfere with the recording. Several different techniques can be used to distinguish the true neural response from the CM. First, the CM occurs much earlier in time than the FFR because it is a pre-neural response. Therefore, a “response” that occurs before 6-10 ms can be attributed to the CM (Skoe & Kraus, 2010). For a further explanation on distinguishing the CM from the FFR refer to the “History” section of this literature review. Secondly, the CM is not affected by stimulus rate whereas the FFR can only be recorded up to a certain stimulus rate before the response begins to degrade (Skoe & Kraus, 2010). A third technique is to simultaneously present a masking noise while recording the FFR. The masking eliminates the FFR, but the CM is unaffected (Aiken & Picton, 2008). Aiken and Picton (2008) ensured that their recordings were true neural responses by presenting a masking noise at 50 dB

HL. A present response with the masking noise on would indicate that the response was the CM and not a true neural response (Aiken & Picton, 2008).

Myogenic or muscle activity can contaminate the desired neural response and can be controlled for by eliminating the myogenic response from the final averaged neural response (Skoe & Kraus, 2010). Myogenic contamination can be greatly reduced if the subject is quiet and resting during the recording. However, artifact rejection based on amplitudes may be used to help eliminate the myogenic responses. The myogenic interference is frequently larger in amplitude than the neural response. For example, Russo et al. (2004) rejected responses with greater than 35  $\mu\text{V}$  amplitudes, as these responses were attributed to eye blinks. Therefore, the artifact rejection criteria should be set at specific amplitude values so that the high amplitude muscle artifact is automatically rejected during the final averaging. For the purpose of this study, the artifact rejection will be set from  $\pm 20 \mu\text{V}$  to  $\pm 75 \mu\text{V}$ .

### **Subject Parameters**

The following subject parameters are also important to consider when recording the FFR: subject state and age. The current study will be conducted on normal hearing adults and therefore the age of the subject will not be relevant for the present study.

### ***Subject State***

Research has shown that the FFR is best recorded in a passive state, either asleep or relaxed (Skoe & Kraus, 2010). It is advantageous during the recording that the individual relax and close his or her eyes to help reduce the level of the overall background noise. For the

purpose of this study individuals will be instructed to relax and sleep if possible. An active state, such as attentively listening to the stimulus, is not recommended during the FFR recording because the literature has shown that this alters the recording. For example, Galbraith and Doan (1995) have observed the effects of attention on FFR amplitudes. Galbraith and colleagues recorded the FFR on 32 subjects using a 400 Hz pure tone and a tone with a missing fundamental (MF). The tones were presented to both ears in a variety of different sequences and the subjects were told to attend to the right or the left ear (Galbraith & Doan, 1995). Subjects were asked to listen for the “target stimulus” which was greater in intensity or longer in duration than the other stimuli. These researchers found that there was a significant increase in FFR amplitudes for both types of stimuli (tone and MF) when the subject was attending to the ear and target stimuli (Galbraith & Doan, 1995). Similarly, Lehman and Schonwiesner (2014) found that the FFR amplitude is affected when the subject’s auditory attention is directed towards a target stimulus as compared to a subject who is passively listening to a stimulus. Specifically two different stimuli, /i/ and /a/, with different fundamental frequencies, were presented in a dichotic and diotic listening situation. A target stimulus was a stimulus that was slightly different in frequency from the /i/ and /a/ stimuli and was randomly presented within the presentation of the regular stimuli. Participants were asked to selectively attend to the target stimulus and indicate when he or she heard the target stimulus. Overall Lehman and his colleague stated that there was a “10 dB” increase in signal amplitude at the fundamental frequencies of the vowels in both listening scenarios. The findings from these two studies support the notion that attention to the stimuli used to elicit the FFR can modulate the amplitude of the neural response. Therefore, in order to ensure higher levels of the system are not influencing this subcortical response, subjects for this study will be asked to remain passive throughout the FFR recording.

Furthermore, functional MRIs have shown that when the individual actively attends to the stimulus, “activation of the subcortical structures” is altered (Skoe & Kraus, 2010). Rinne, Balk, Koistinen, Autti, Alho, and Sams (2008) investigated the effects of auditory spatial selective attention on neural encoding. Rinne and colleagues observed these effects using the fMRI. Iterated rippled noise with different pitches (115 Hz, 190 Hz, 800 Hz) was simultaneously presented to the right and left ear. Subjects were asked to attend to one ear and indicate by pushing a button when they perceived a change in stimulus pitch. The fMRI suggested that the auditory cortex and inferior colliculus contralateral to the attended ear exhibited a higher level of activity than the ipsilateral IC or cortex. The inferior colliculus is a known neural generator of the FFR. Therefore active listening will be discouraged during this study to avoid potential interference from higher brain structures that could alter the true subcortical response.

### ***Subject Age***

Studies have shown that neural phase locking abilities change with increasing age (Clinard, Tremblay, & Krishnan, 2010). Clinard and colleagues studied the effects of aging on the brainstem’s ability to phase lock to several different pure tones. The twenty eight subjects for Clinard and colleague’s study ranged from age 22 years to 77 years. Results showed that age did not affect the brainstem’s ability to follow the frequency of a 463 Hz, 498 Hz, and 500 Hz pure tone stimulus. However, age negatively affected the brainstem’s ability to phase lock to a 925 Hz, 998 Hz, and 1000 Hz pure tone stimulus. Clinard and colleagues found that the neural responses to the 1000 Hz stimulus were not as robust as the neural responses to the 500 Hz stimulus. The overall conclusion was that age will reduce the effectiveness of the brainstem’s ability to follow the frequency of the stimulus when the stimulus is 1000 Hz or above. Overall this study shows that, in general, age can affect how well the brainstem can phase lock to certain

frequency stimuli. This study sheds light on the impact of age on the brainstem's ability to phase lock to lower frequencies. In the current study used an /u/ stimulus with a fundamental frequency of 120 Hz. Therefore, based on the results from the previous study, age would not affect the brainstem's ability to encode this low frequency fundamental frequency. However, young adults ranging in age from 20 to 25 years of age will be used for this study. Therefore the effects of aging on the FFR will not be a concern. The following section describes the effects of background noise on overall speech acoustics and vowel perception, since the current study will use different SNR values to determine how noise impacts the FFR.

### **Effects of Background Noise on Speech Acoustics**

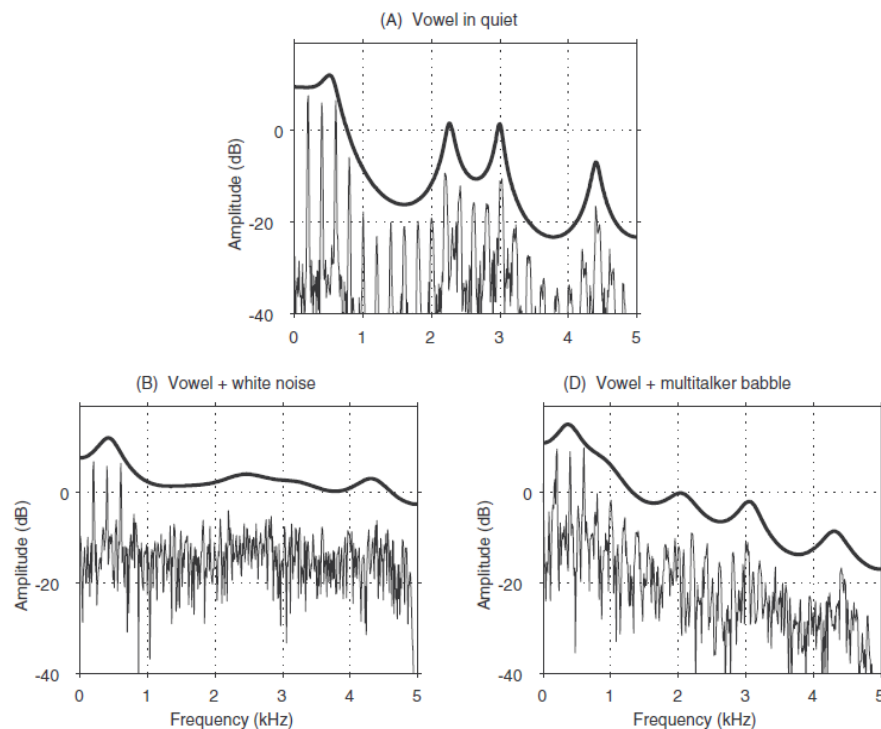
Background noise is present in everyday listening situations, and decreased speech intelligibility in noise is a common complaint among individuals with both normal hearing and hearing loss (Assmann & Summerfield, 2004). Investigators have used different types of noise to study the effect of background noise on speech intelligibility (Assmann & Summerfield, 2004). Examples of noise stimuli used to simulate the effects of everyday noise on speech perception are narrowband noise, pure tones, broad band noise, and multitalker babble or competing speech (Assmann & Summerfield, 2004). All types of noise stimuli can be defined as “unwanted sound containing major components within a certain frequency range” (Berglund, Hassmen, & Job, 1996, p. 2986). When a noise stimulus is presented concurrently with a speech signal, acoustical characteristics of the noise signal, such as its intensity and spectrum, can alter the acoustical make-up of the speech signal, thereby leading to poor speech perception. These factors will be discussed further in the next section.

Studies have shown that one of the factors that significantly affects speech understanding in noise is the intensity of the noise in relation to the intensity of the speech signal (Katz, 1994).

The intensity of the desired speech signal compared to the intensity of the background noise is known as the Signal to Noise ratio (SNR). A high SNR indicates that the speech signal is louder than the background noise. For example, if the desired speech signal is 60 dB and the background noise is 50 dB, the SNR is + 10 dB, meaning that the speech is 10 dB above the background noise (Katz, 1994). A negative SNR indicates that the background noise is louder than the desired speech signal. When the background noise becomes louder than the desired speech signal the speech is “masked” (Katz, 1994). A high level of background noise will affect the entire speech waveform, including the consonants and vowels of a word (Drullman, 1995). Consonants do not carry as much energy as do vowels and are therefore more susceptible to noise masking. Studies have been conducted to investigate the effects of unfavorable SNRs on both consonant and vowel perception (Miller & Nicely, 1955; Pickett, 1957). The incorrect identification of a consonant is primarily due to a misperception of the manner of articulation, such as a “p” being perceived as a voiced consonant, “b” (Katz, 1994). However, the incorrect identification of a vowel is primarily due to misperception of the vowel formant frequencies (Katz, 1994). The remainder of this section will focus on the effects of noise on the acoustics and perception of vowels since a vowel stimulus will be used to evoke the neural response in the current study.

A listener identifies a vowel by utilizing the acoustical information present in its first three formants. As described in the basic acoustics section, formants are the peaks of energy located at specific frequencies, representing acoustic resonances of the vocal tract during vowel production. Assmann and Summerfield (2004) investigated the effects of noise on the formant peaks of the vowel /e/ in the word /head/. In this study, Assmann and Summerfield (2004) compared the FFTs of vowel stimuli in quiet and in the presence of white noise and multi talker

babble at 0 dB SNR. A FFT of the vowel /e/ in quiet (figure A), white noise (figure B), and multitalker babble (figure D) is pictured below. Results indicated that the energy present at the formant frequencies of the vowel was reduced in amplitude and there were “additional spectral peaks” present in the vowel spectrum in the presence of multi-talker babble as compared to the “clean” (quiet, no noise) condition (Assmann & Summerfield, 2004). In the white noise test condition, the formant frequencies F2, F3, and F4 are no longer present in the stimulus spectrum. Formant representation in the stimulus spectrum of the vowel was significantly affected in the presence of both kinds of noise stimuli. Thus, the findings from Assmann and Summerfield (2004) indicate that concurrently presented background noise can alter the acoustical characteristics of the speech signal.



*Figure 22.* FFTs of vowel /e/ in quiet, in the presence of white noise, and in the presence of multi talker babble at 0 dB SNR



Leek and Summers (1996) similarly stated that there is a loss in the contrast between the first three formants of a vowel /i/ when comparing the spectrogram in quiet versus noise.

Overall, changes in the acoustic make-up of the vowel signal in the presence of background noise leads to difficulties in vowel perception.

Studies have also shown that the frequency or average spectrum of the noise will also influence the degree to which the noise “masks” the speech signal and hence speech perception in noise. The most effective noise masker is one that contains energy across the frequencies that are present in the long term average spectrum of the desired speech signal (Berglund et al, 1996; Katz, 1994). The long term average speech spectrum is a representation of the “intensity of the speech as a function of frequency” (French & Steinberg, 1947, p. 3). Speech energy typically falls within 100 Hz to 6000 Hz and therefore noise falling within this range is an effective masker (Berglund et al., 1996). Additionally, low frequency noise creates an “upward spread of masking” effect at high presentation levels. As the intensity of the low frequency noise increases, the range of frequencies in the stimulus that are masked by the noise also increases, particularly the higher frequencies of the stimulus (Gelfand, 2009). The upward spread of masking phenomenon occurs because the high frequency regions of the basilar membrane are stimulated or excited by this low frequency sound which “masks” the higher frequency stimulation of the basilar membrane (Gelfand, 2009). In conclusion, correct identification of a vowel depends both on the intensity and spectrum of noise that is interfering with the desired speech signal. The current study used varying SNRs to investigate the effect of noise on vowel perception, and therefore the intensity of the noise on vowel perception will be discussed further in the remaining section.

### *Vowel perception in noise*

Multiple studies have been conducted to measure the impact of the intensity and spectra of background noise on vowel perception and identification. In order to study the effects of the intensity of background noise on vowel perception, Leek, Dorman, and Summerfield (1987) investigated the ability of 20 normal hearing listeners to correctly identify four different digitally created stimuli that approximated the vowels (/i/, /æ/, /u/, /ɑ/). The four stimuli were presented at 85 dB SPL through headphones, in the presence of simultaneous contralateral noise stimulation. Leek et al. (1987) reported that normal hearing individuals needed approximately a 1- 2 dB difference between the spectral peaks at the formant frequencies of the vowel and the “trough”, i.e. the rest of the stimulus spectrum, to correctly identify a vowel presented in quiet. However, when background noise was presented contralaterally, a 4-dB difference between the formant peaks and trough (the rest of the stimulus spectrum) was needed for accurate identification of the vowel. Vowel identification in background noise was poor if the “peak to trough” difference was less than 4 dB. Based on these findings, Leek et al. (1987) suggested that noise reduces the difference between the peaks and troughs of the vowel spectrum, leading to poorer understanding of vowels in noise.

Similarly, Leek and Summers (1996) investigated the differences between the peak formant frequencies and the trough of vowel like stimuli in varying levels of broadband noise and its impact on vowel perception. Four stimuli were digitally created so that first and second formant frequencies of the stimuli resembled that of the first and second formant frequencies of the vowels /æ/, /ε/, /ɪ/, and /i/. Six participants with normal hearing and six participants with hearing loss were asked to identify which vowel he or she perceived when the noise was

presented at 75 dB SPL, 80 dB SPL, and 85 dB SPL. Within each of the three listening conditions (noise presented at 75 dB SPL, 80 dB SPL, and 85 dB SPL) the spectral contrast between the peak of the stimulus (formant frequencies) and the trough of the vowel of the stimulus was increased in 2 dB increments from 2 dB to 24 dB. Results indicated that as the spectral contrast increased from 2 to 24 dB the percentage of vowels correctly identified by all normal hearing and impaired listeners increased. However, the highest noise level (85 dB SPL) resulted in the lowest number of vowels correctly identified across all levels of spectral contrast (2 dB o 24 dB) for both normal hearing and impaired hearing individuals. Similar to Leek et al. (1987), this part of the investigation indicates that “formant” regions of the vowel-like stimulus are not as pronounced in increasingly noisy listening environments. The decrease in spectral contrast is due to a “smearing” effect of the spectral properties of the stimulus and therefore a greater difference between the spectral peaks is necessary for correct vowel perception in noise.

Another factor that affected how well normal hearing and hearing impaired individuals were able to discriminate vowels in this presence of noise in this study was the similarity between F1 and F2 frequency regions of the four stimuli. The frequency region where F1 and F2 were present in each of the four stimuli is listed below in Table 1. A comparison between F1 and F2 of the four vowel-like stimuli shows that some vowels have similar F1 and F2 frequency regions, while other vowels do not have similar formant frequency regions. For example, as seen in the table below the vowels /æ/ and /i/ are “acoustically distinct” in that the frequency region for F1 and F2 is completely different. On the other hand, /ɪ / and /i/ have identical F1 frequency regions and similar F2 frequency regions. A comparison between /æ/ and /ɛ/ shows that these vowels have the same F2 frequencies and different F1 frequencies, whereas the vowels /ɪ/ and /i/ have the same F1 frequencies and different F2 frequencies.

Table 2

*F1 and F2 formant frequency regions for vowel like stimuli /æ/, /ε/, /ɪ/, and /i/*

Vowel	F1 (Hz)	F2 (Hz)
/æ/	600, 700	1800,1900
/ε/	400, 500	1800, 1900
/ɪ/	300, 400	1900, 2000
/i/	300, 400	2100,2200

For normal hearing individuals, vowels with less similar formant frequency regions, such as /æ/ and /i/, were more easily identified; whereas vowels with overlapping formant frequency regions such as /ɪ/ and /i/ were not as easily differentiated. Leek and Summers (1996) found that the hearing impaired individuals relied on the F1 to identify the correct vowel and were unable to utilize F2 as a characteristic for vowel discrimination. Therefore, vowels with similar F1 frequency regions, such as /ε/ and /i/, /ε/ and /ɪ/, and /ɪ/ and /i/ were extremely difficult to distinguish for hearing impaired listeners. Both normal hearing and hearing impaired individuals struggled more to identify all vowels, regardless of overlapping formant frequency regions, with increasing noise intensity. However, hearing impaired individuals did not identify vowels as well as the normal hearing individuals when noise levels were high (85 dB SPL). These results indicate that the ability to differentiate between F1 and F2 is important for correct vowel

perception. Increasing noise intensity combined with hearing loss, negatively impacts an individual's ability to perceive a vowel based on F1 and F2 frequencies.

Similar to Leek and Summers (1996) additional studies have been conducted to determine the spectral cues important for listeners to correctly identify vowels in noise (Nabalek & Dagenais, 1986; Nabalek, 1988). Nabalek and Dagenais (1986) initially investigated the ability of ten, middle aged hearing impaired listeners to understand fifteen monophthongs and diphthongs in a quiet condition versus a noisy condition (0 dB SNR). Prior to performing this study, it was determined that normal hearing individuals could easily identify all vowels in both listening conditions. Stimuli were presented via TDH-50 headphones, and participants were asked to repeat out loud the vowel he or she perceived. A monophthong is a pure vowel that is produced with the articulators in one, fixed position, whereas a diphthong is two vowel sounds combined together. Results indicated that the most common error hearing impaired listeners made in noise was misidentifying monophthongs when the formant frequencies between monophthongs were similar. Monophthongs which were only discriminable by one formant frequency were easily misidentified when the frequency necessary for identification was masked by the noise. Nabalek and Dagenais (1986) found that monophthongs commonly confused in the noise condition were /ɛ/ -- /i/, /ɜ/ -- /ʊ/, and /ɑ/ -- /ʌ/. Diphthongs were incorrectly identified as monophthongs when the hearing impaired individual only perceived the first monophthong of the vowel. For instance, /aɪ/ (pronounced in the word "sky") was perceived as the monophthong /ʌ/ (pronounced in the word "up"). This misperception may be explained by a masking of the second monophthong, where a change in formant frequencies is not detected and therefore only the first monophthong is heard. Nabalek (1988) expanded on the findings of the previous study and examined the impact of various degrees of hearing loss and varying ages (ranging from 20 to 80 years of age) on vowel

identification in noise. Results from this study confirmed that the common vowel misperceptions in noise occurred between monophthongs with similar F1, F2, and F3 frequencies. Individuals with more severe hearing losses confused monophthongs more easily, since the differences between formant frequencies needed to be greater for correct vowel identification. Furthermore, Nabalek (1988) confirmed that first part of the diphthongs, perceived as monophthongs, consisted of formant frequencies that were similar to the formant frequencies of monophthongs. In conclusion, high noise levels and hearing loss can cause listeners to misinterpret a vowel due to less distinct formant frequencies. For the purpose of this study, all participants had normal hearing and therefore hearing loss is not a contributing factor for decreased vowel perception.

Studies have also examined the acoustic and perceptual effects of different noise spectra on speech signals as well. In 1957, Pickett sought to determine the effects of flat, low frequency, and high frequency broadband noise on vowel acoustics and identification. Seven normal-hearing listeners were asked to identify different vowels presented in the context of syllables in quiet and in the presence of noise. The noise was digitally mixed with the syllables and presented to the individuals via headphones and loudspeakers. It was determined that low frequency noise tended to mask the first formant frequency information, while flat and high frequency noises disrupted the second formant more. As mentioned earlier, both first and second formant frequency information is critical for vowel identification. The second formant is necessary for identifying certain vowels which have similar first formant frequencies, whereas the first formant is necessary for identifying vowels with similar second formant frequencies. Therefore when these critical first and second formants are not available to the listener, vowels with similar first formants can become difficult to differentiate. Hence, these results indicate that speech

intelligibility of various vowel stimuli is also significantly affected by the frequency content of the competing noise.

### **Neural Encoding in the Presence of Background Noise**

Anderson and colleagues have suggested that neural responses, such as the ABR and FFR, can be recorded in the presence of background noise to investigate the effects of adverse listening conditions on the brainstem's ability to accurately process speech stimuli. A comparison of the results obtained for subjects' performance on behavioral speech identification tasks as well as their objective neural responses to speech stimuli can help to identify what brain processes are associated with "favorable" speech perception (Anderson et al., 2010).

Speech perception requires that the brain preserve the "temporal characteristics" of the speech signal as it travels along the auditory pathway. Neural synchrony is essential in "preserving the time cues needed for speech perception" (Anderson et al., 2010). The FFR preserves these temporal characteristic of the speech signal such as the envelope (F0) and TFS (formant) cues via neural phase locking and is therefore an effective electrophysiologic tool to study neural encoding of speech in noise.

A near field recording was conducted by Henry and Heinz (2012) to investigate the effects of broadband noise on the neural synchrony and temporal encoding of tones. This study investigated electrophysiological responses in chinchillas rather than humans. Electrodes were placed on single auditory nerves and the threshold of each fiber was recorded along with the characteristic frequency for the auditory nerve. Noise was then digitally mixed with the tone and was presented at 10, 15, and 20 dB above the tone. Henry and his colleague discovered that the thresholds of the auditory fibers increased with increasing levels of noise. Additionally, the fibers

were not as sharply tuned to a specific frequency when the noise was present. Henry and Heinz concluded that this “broader tuning curve” of the auditory nerve fiber in the presence of noise most likely would result in reduced temporal coding.

Anderson and colleagues also hypothesized that a noisy environment may reduce neural synchrony, which in turn may lead to decreased speech intelligibility (Anderson et al., 2010). In order to examine the effects of background noise on neural synchrony of speech encoding, Anderson and colleagues recorded the FFR in sixty children using a vertical electrode montage (Anderson et al., 2010). The responses were evoked using a /da/ stimulus presented at 80 dB SPL in quiet as well as in the presence of multi talker babble set at an SRN of +10 dB. The FFR was evaluated by analyzing the latencies of the response that were “phase locked” to the formant frequencies of the syllable. The results on this electrophysiological response for each child were then compared to their behavioral results on the Hearing in noise test (HINT). Results from the HINT lead to categorizing the children into two separate groups: the children who scored in the 50<sup>th</sup> percentile or above on the HINT were categorized as group 1 and those who scored below the 50<sup>th</sup> percentile were placed in group 2. Children were also assessed based on their reading level and placed into two groups based on their performance on the Test of Word Reading Efficiency-Total (TOWRE-T). Anderson and colleagues reported that the latencies of the neural response were delayed in all of the children when noise was presented during the recording. Specifically, the phase locked response to the formant frequencies was delayed. Additionally, these investigators reported that the children in group 2, who scored below the 50<sup>th</sup> percentile on the HINT and who were poor readers, exhibited greater latency shifts in FFR in comparison to the children in group 1. Mean latency data for group 1 and group 2 at certain peak latencies are



provided in Table 2 below. These mean latencies values were then compared to the latencies of the stimulus waveform with peaks occurring at 32 ms, 42 ms, and 52 ms

Table 3

*Mean latency FFR values for children with HINT scores in the top and bottom 50<sup>th</sup> percentile and the mean latencies for the top and bottom readers*

Peak	HINT (top 50 <sup>th</sup> )	HINT (bot. 50 <sup>th</sup> )	Top readers	Bottom readers
34 ms	35.32 ms	35.39 ms	35.11 ms	35.53 ms
42 ms	43.53 ms	43.89 ms	43.55 ms	43.83 ms
52 ms	53.45 ms	53.57 ms	53.21 ms	53.76 ms

Mean latency FFR values for children with HINT scores in the top and bottom 50<sup>th</sup> percentile and the mean latencies for the top and bottom readers. A strong correlation was found between the neural synchrony and HINT scores and reading ability. Specifically, Anderson and colleagues concluded that there is a correlation between poor perception of speech in noise (as seen by performance on the HINT) and “decreased neural synchrony” as reflected in the delayed latencies of the FFR. These results support that neural encoding is disrupted by noise.

It is clear from these behavioral and neural studies that background noise impairs one’s ability to perceive and encode speech optimally. Difficulty understanding speech in noise is a common complaint from individuals with sensorineural hearing loss. A sound understanding of neural encoding patterns in background noise in normal hearing human listeners is necessary in order to understand the underlying neurophysiology in listeners with SNHL. There are only a

few published studies investigating neural encoding of speech in noise in normal hearing human listeners. Hence, the purpose of this study is to use the FFR to further investigate the effects of noise on the neural representation of vowel sounds in normal hearing individuals.

### **CHAPTER 3: METHODS**

#### ***Participants***

Six young adults participated in the study (mean age = 23.83 years, age range: 22--25, 3 male; 3 female). All participants were required to have normal hearing sensitivity. Normal hearing sensitivity was defined as air conduction thresholds better/lower than 15 dB HL at octave frequencies from 250 Hz- 8000 Hz in each ear. Tympanometric measures indicated normal middle ear function for all participants (Peak pressure range = -150 to 50 daPa; Compliance = .3 to 1.4 ml; ear canal volume = .6 to 1.5 ml) (Roeser & Valente, 2007). Self-reported otologic and neurologic histories were unremarkable. Participation in this study was voluntary and all participants were required to give their informed consent before participating in this study (Appendix A). Prior to testing, all procedures were approved by the Institutional Review Board of Towson University (Appendix B).

#### ***Stimuli***

A synthetically produced steady-state vowel stimulus /u/ (F0=120 Hz, F1=360 Hz) was used to evoke the FFR. Etymotic Research ER-3A insert headphones were used for monaural stimulation of the right ear and the stimulus was delivered at an intensity level of 75 dB SPL. The total stimulus duration was 260 ms and delivered at a rate of 3.13/second. An alternating stimulus polarity (condensation and rarefaction) was used during all FFR recordings in order to facilitate teasing apart of the envelope and TFS components of the FFR. Background noise was presented simultaneously with the vowel /u/ in three different listening conditions with varying

signal to noise ratios (SNRs). Speech shaped noise was used and is noise which represents the long term average speech spectrum. The FFR was recorded in the following listening conditions: + 5 SNR, 0 SNR, and -5 SNR. The +5 SNR means that the stimulus is 5 dB above the noise, 0 SNR means that the signal is the same level as the noise, and -5 means that that the signal is 5 dB less than the noise. The noise and stimulus was presented to the same ear but routed via two different channels of the IHS system.

### ***Procedure for FFR data collection***

During the FFR recording, participants were seated in a comfortable recliner in a sound treated booth. Participants were asked to relax with their eyes closed, and were strongly encouraged to sleep in order to minimize body movements that could contribute to artifacts in the response. The duration of the testing session was approximately three hours. The FFR was recorded using a /u/ stimulus presented at an intensity of 75 dB SPL and a stimulus rate of 3.13/s at various SNRs. The FFR was collected using a two-channel recording using the Intelligent Hearing System SmartEP equipment. A vertical electrode montage was used to record the FFR, consisting-of a non-inverting (+) electrode placed at Cz (vertex), two inverting electrodes (-) at M2 (right mastoid) and C7 (nape of the neck), and a ground electrode placed at Fpz (forehead). The duration of the averaging window was set at 300 ms. The sampling rate used was 10,000Hz. 2000 sweeps were used to obtain individual waveforms. Since the majority of the energy found within the FFR is lower to mid frequencies, the analog band-pass filter was set at 100-3000 Hz. The artifact rejection level was set between  $\pm 20 \mu\text{V}$  to  $\pm 75 \mu\text{V}$ , so that amplitudes obtained outside of this range were not included in the final average.

### ***FFR: Data Analysis***

FFR data were exported from IHS to MATLAB for analysis in the temporal and frequency domains. Temporal waveforms were obtained for each subject per condition. A separate grand average waveform was calculated for the response to the rarefaction stimulus and the response to the condensation stimulus for each listening condition. Both qualitative and quantitative data analysis techniques were used to examine the FFR data.

#### *Qualitative data analysis:*

Qualitative indices of the neural response at each SNR was provided through a visual analysis of the periodicity and root mean square (RMS) amplitude of the grand averaged waveforms obtained at each SNR.

#### *Quantitative data analysis:*

MATLAB software was used to analyze the temporal waveforms in the frequency domain. The FFR is a complex response and the FFT analysis was used to break down the response into the constituent frequencies. Specifically, an FFT analysis was conducted on the temporal response waveforms obtained for each subject for every condition to determine the distribution and magnitude of energy in the FFR spectrum. As the FFR represents neural phase locking to the envelope (F0) and TFS (F1) of the stimulus, we expected to see maximal energy at peaks in the FFR spectrum corresponding to the F0 (120 Hz) and F1 (360 Hz) of the stimulus. The magnitude of the response peaks at 120 Hz and 360 Hz were recorded across all subjects and condition.

## Statistics

Descriptive statistics were used to analyze the data instead of inferential statistics due to the small number of participants in this study. The FFT analysis provided the amplitude ( $\mu\text{V}$ ) of the response at specific frequencies (120 Hz and 360 Hz). Mean response amplitude values were then calculated for all six participants in the clean stimulus condition, +5 SNR, 0 SNR, and -5 SNR condition. Standard deviations were also obtained. Standard deviations provided information about how much each individual frequency varies from the mean amplitude of the given frequency. Trends were evident in the data such as a decrease in amplitude at specific frequencies with increasing noise levels.

While such descriptive statistical methods will not allow us to test for differences across conditions, they will shed light on emerging trends in the data. The chart below is an example of how data will be organized.

Table 4

*Sample data table: Amplitude ( $\mu\text{V}$ ) of the response at 120 Hz determined by the FFT analysis*

	Clean Stimulus	SNR: + 5 dB	SNR: 0 dB	SNR: -5 dB
Participant 1	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$
Participant 2	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$
Participant 3	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$
Participant 4	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$
Participant 5	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$
Mean amplitude	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$
Standard Deviation	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$	___ $\mu\text{V}$

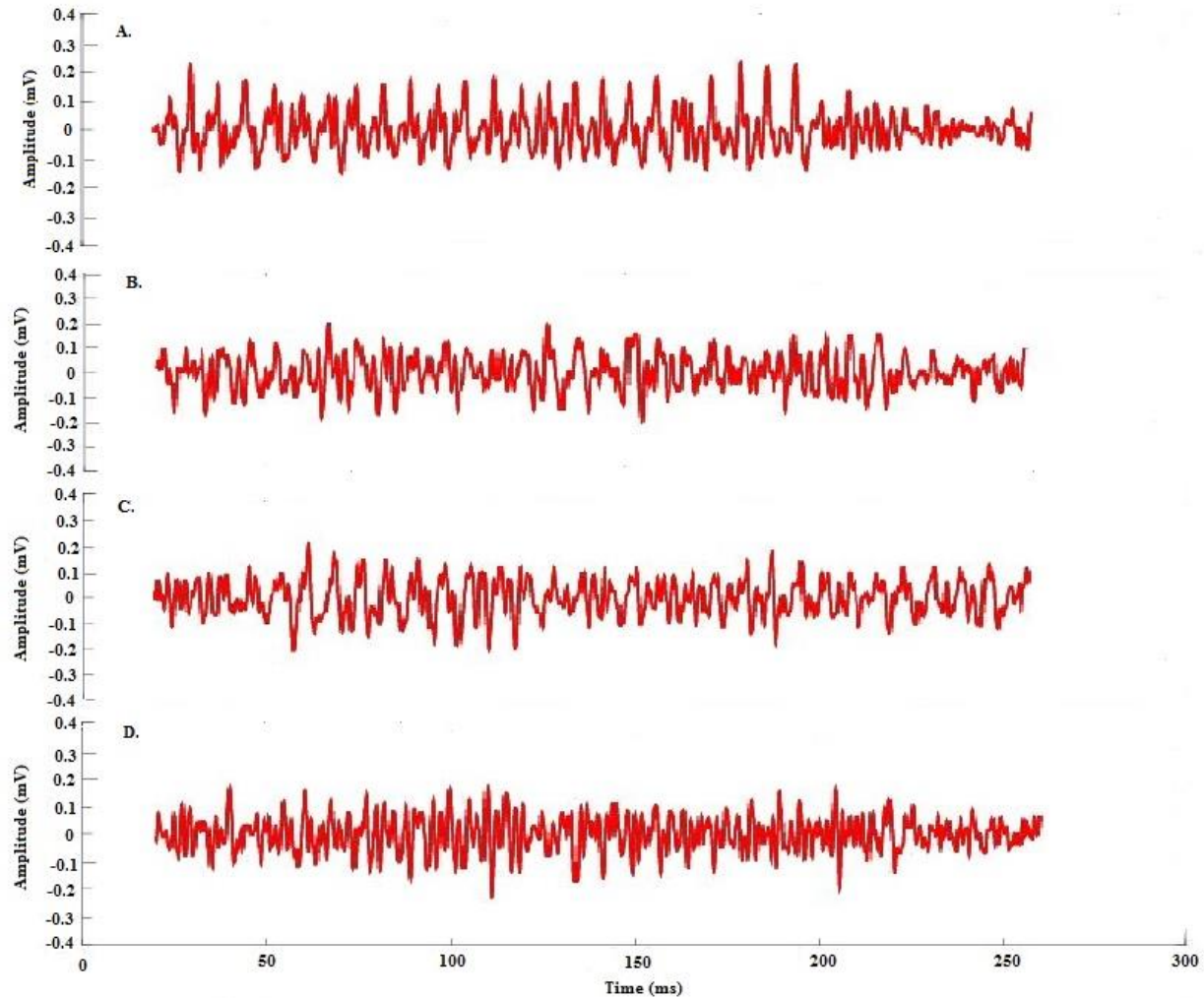
## **CHAPTER 4: RESULTS**

The results section will be divided into two primary sections focusing on brainstem neural encoding to the fundamental frequency (F0) of the stimulus, or envelope encoding ( $\text{FFR}_{\text{ENV}}$ ) and brainstem neural encoding to the first formant (F1) frequency of the stimulus, or temporal fine structure encoding ( $\text{FFR}_{\text{TFS}}$ ) respectively. Within each of these primary sections, results from temporal grand averaged temporal waveforms and spectral analyses (FFT analysis of F0 and F1 (first formant and fundamental frequency) and grand averaged FFT data) will be discussed for each of the four listening conditions (clean, +5 SNR, 0 SNR, -5 SNR).

### **$\text{FFR}_{\text{ENV}}$ responses related to the Fundamental Frequency (F0)**

#### ***Grand Average FFR Temporal waveforms***

The grand average temporal waveforms (averaged across six participants) of the summed envelope FFR ( $\text{FFR}_{\text{ENV}}$ ) are displayed below in Figure 23 for each test condition. Each grand average temporal waveform obtained for the four background noise conditions is represented in a separate panel. Panel A represents the grand average temporal waveform of the  $\text{FFR}_{\text{ENV}}$  in the clean background noise condition (/u/ stimulus with no noise present). Panel B represents the grand average temporal waveform of the  $\text{FFR}_{\text{ENV}}$  for the + 5 SNR background noise condition. Panel C represents the grand average temporal waveform of the  $\text{FFR}_{\text{ENV}}$  for the 0 SNR background noise condition. Lastly, Panel D represents the grand average temporal waveform of the  $\text{FFR}_{\text{ENV}}$  for the -5 SNR background noise condition. This organization of the panels will be followed in all subsequent figures in the results section.



*Figure 23.* Grand average temporal waveforms for the summed envelope FFR ( $\text{FFR}_{\text{ENV}}$ ) for the four background noise conditions (Panel A: Clean condition, Panel B: +5 SNR condition, Panel C: 0 SNR condition, Panel D: -5 SNR conditions).

Visual inspection of  $\text{FFR}_{\text{ENV}}$  waveforms in Figure 23 reveals several interesting trends related to the mean amplitude values and the periodicity of the grand average temporal waveforms. First, the overall amplitude of the grand average temporal waveform substantially decreases as the severity of the condition worsens (+ 5 SNR background condition in Panel A to - 5 SNR background noise condition in Panel D). Specifically, the mean amplitude of the grand average temporal waveform is approximately .2 mV in the clean condition, .15 mV in the +5 dB

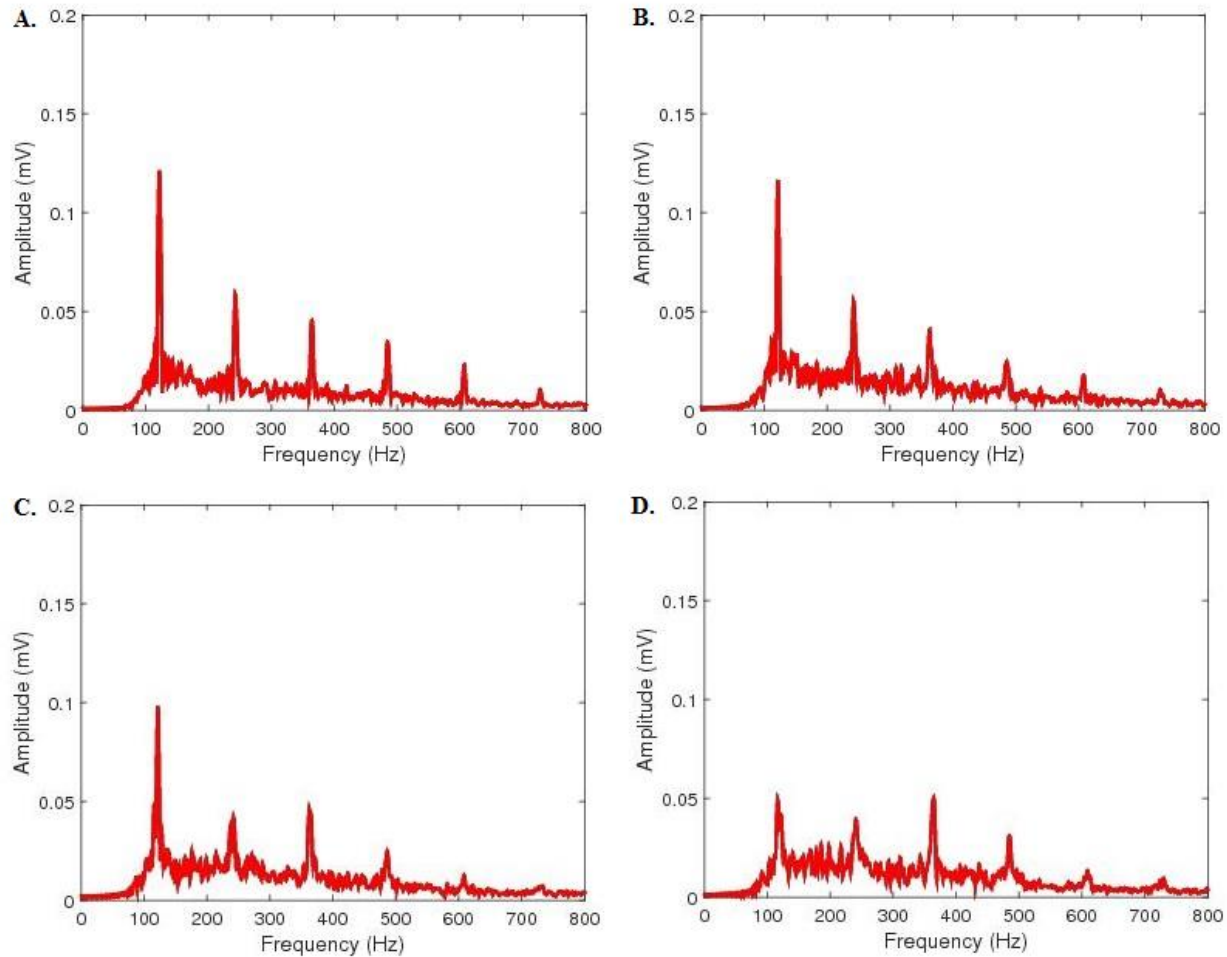
SNR condition, and .1 mV in the 0 and -5 dB SNR background noise conditions. The difference in mean amplitude of the grand average temporal waveform from the clean condition to the -5 SNR condition is approximately a 50 % decrease.

Second, the overall morphology of the grand average temporal waveform declines as background noise increases, as reflected in temporal resolution in each condition. Specifically, the temporal resolution of the waveform is most clearly defined in the clean condition, where the greatest amount of periodicity can be observed. In other words, defined peaks occurring at discrete time intervals throughout the duration of the response (250 ms), are noted in the clean condition. As the listening condition degrades, the grand average temporal waveform loses the temporal resolution or periodicity that was present during the clean condition. The -5 dB SNR background noise condition exhibits the poorest temporal resolution out of all four listening conditions throughout the duration of the stimulus, exhibited in the decreased periodicity, or loss of defined regularly spaced peaks over the response duration.

***Spectral Analysis: Fast Fourier Transform (FFT) of the  $FFR_{ENV}$***

FFT analysis was then used to convert the  $FFR_{ENV}$  grand average temporal waveforms into the frequency domain for spectral analysis across all four background noise conditions and is represented in Figure 24 below.





*Figure 24.*  $FFR_{ENV}$  grand average FFT analysis for each of the four background noise conditions (Panel A: Clean condition, Panel B: +5 SNR condition, Panel C: 0 SNR condition, Panel D: -5 SNR condition). Amplitude of the response is displayed along the y-axis and frequency is displayed along the x-axis

Visual inspection of the grand average FFT for each background noise condition indicates several interesting trends in the spectral data. First, as seen in Figure 24, the amplitude of energy present in the response at the F0 (120 Hz) decreases as the severity of the condition worsens. Specifically, the amplitude at the F0 is approximately .12 mV for both the clean and + 5 dB SNR listening conditions. The F0 amplitude then decreases to approximately .1 mV and .05 mV for the 0 and -5 dB SNR conditions, respectively. As noted by these approximate values, while the

absolute F0 magnitude appears to be generally similar among the clean, + 5 dB SNR, and 0 dB SNR background noise conditions, the energy present at the F0 in the -5 dB SNR background noise condition is substantially lower than the F0 amplitude in any other listening condition.

Secondly, in all of the background noise conditions, the magnitude at the F0 is considerably greater than the magnitude at the harmonics (whole number multiples of the F0 such as 240 Hz, 360 Hz, 480 Hz, etc.). This pattern is true in all of the background noise conditions, with the exception of the -5 dB SNR background noise condition. Specifically, the energy present at the harmonics is at least 50 % lesser than the energy present at the F0 in the clean, +5 dB SNR, and 0 dB SNR background noise condition. In the -5 SNR listening condition the peak at the F0 is not as pronounced and it is more similar in magnitude to the energy present at the other harmonic frequencies.

### ***Quantitative Spectral Analysis: Mean Values and Individual Data***

Quantitative measures of the brainstem envelope encoding were obtained for each of the six participants by measuring the magnitude at the F0 (120 Hz) in the spectra of the individual FFR<sub>ENV</sub>. Table 5 below summarizes the absolute magnitude of energy in millivolts (mV) at the F0 for each participant as well as the mean F0 amplitude values and standard deviation values for each of the four background noise listening conditions. The far right column displays the difference between the F0 amplitude measured in the clean listening condition and the -5 dB SNR background noise condition for the individual subjects and their mean and standard deviation difference value.

Table 5

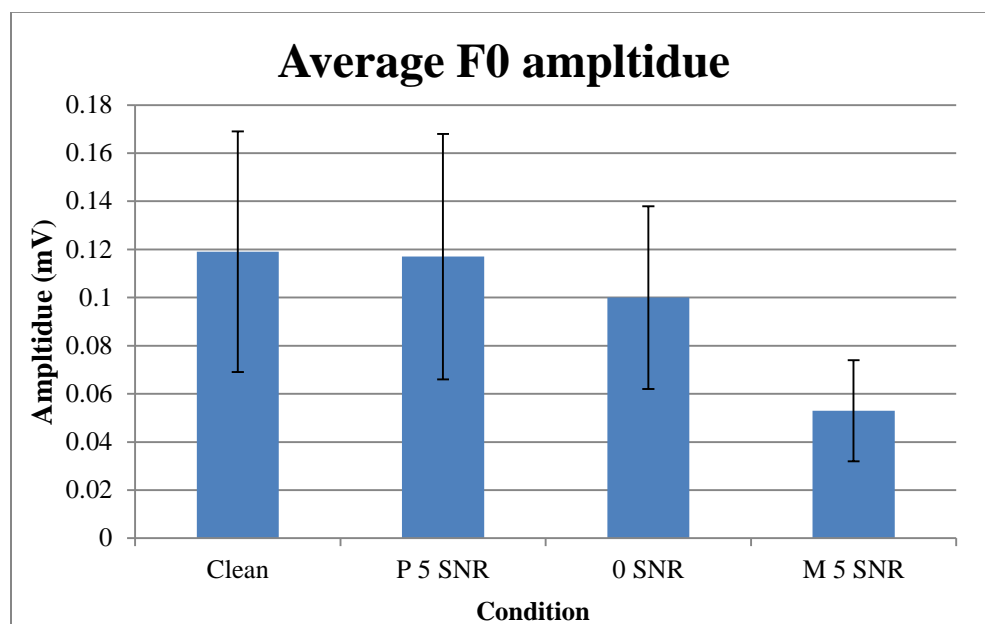
*Individual amplitude measurements at the F0, Mean F0 values, and Standard Deviation*

	Clean	+ 5 SNR	0 SNR	-5 SNR	Difference
Subject 1	0.024 mV	0.028 mV	0.025 mV	0.027 mV	-.003 mV
Subject 2	0.151 mV	0.151 mV	0.126 mV	0.025 mV	.126 mV
Subject 3	0.151 mV	0.119 mV	0.109 mV	0.067 mV	.084 mV
Subject 4	0.117 mV	0.175 mV	0.119 mV	0.065 mV	.052 mV
Subject 5	0.115 mV	0.133 mV	0.116 mV	0.065 mV	.05 mV
Subject 6	0.156 mV	0.098 mV	0.108 mV	0.067 mV	.089 mV
<b>Mean</b>	<b>0.119 mV</b>	<b>0.117 mV</b>	<b>0.100 mV</b>	<b>0.053 mV</b>	<b>.066 mV</b>
<b>SD</b>	<b>0.050 mV</b>	<b>0.051 mV</b>	<b>0.038 mV</b>	<b>0.021 mV</b>	<b>.044 mV</b>

Note. Difference values = Difference between the F0 amplitude in the clean condition and the -5 SNR background noise condition

As seen in Table 5, the mean amplitude values at F0 substantially decreases as the SNR becomes poorer (Clean: 0.119 mV, +5 SNR: 0.117 mV, 0 SNR: 0.100 mV, -5 SNR: 0.053 mV). The most notable trend seen when looking at the mean F0 amplitude values across background noise conditions is the substantial decrease in mean F0 amplitude (.066 mV difference ) occurring between the clean and most severe background noise condition (.119 mV and .053 mV, respectively). On the other hand, the smallest difference in F0 amplitude occurs between the clean and +5 SNR background noise condition (.119 mV and .117 mV, respectively). The variance was calculated for each background noise condition and is represented by the standard

deviation values. In general, the variability was similar ( $\sim .05$  mV) across the first three test conditions (quiet, + 5 dB and 0 dB SNR) and was approximately 50 % lesser (.021 mV) for the - 5 dB SNR condition. This finding reinforces that the amplitude of the F0 was substantially affected by the poorest SNR. The average F0 amplitude for each background noise condition, as well as the variability in each condition can be visually seen in Figure 25 presented below.



*Figure 25.* Average F0 amplitude obtained using FFT analysis of the summed average temporal waveform in the clean, + 5 SNR, 0 SNR, and -5 SNR listening conditions

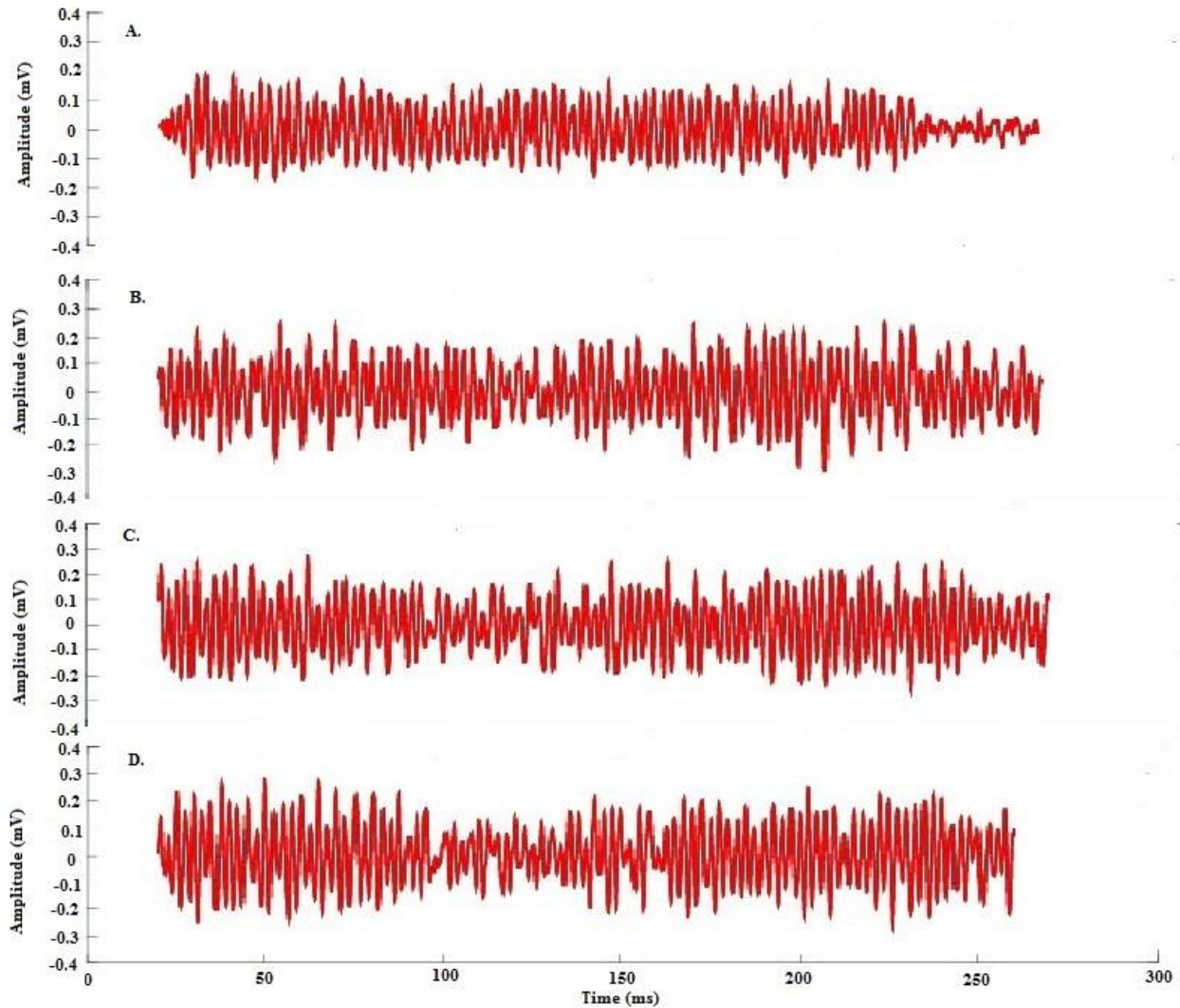
In general the F0 amplitudes for the individual subjects follow the same trend as the mean F0 amplitude for each of the four conditions, as seen in Table 5. First, there is an overall decrease in the F0 amplitude values from the clean to the most severe background noise condition (-5 dB SNR) for 5 of the 6 subjects. The F0 amplitudes decrease from the clean to -5 dB SNR background noise condition by at least .05 mV. These difference values for each subject

are displayed in the far right column of Figure 25 above. The only exception to this general decrease in F0 amplitude as the severity of the condition worsens was participant 1 who shows essentially the same F0 amplitude values across test conditions.

### **FFR<sub>TFS</sub> responses related to the First Formant (F1)**

#### ***Grand Average FFR Temporal waveforms***

The following section will discuss the data collected for the subtracted grand average temporal waveform (FFR<sub>TFS</sub>) which represents subcortical neural encoding of the first formant frequency of the stimulus (360 Hz). Grand average temporal waveforms for the FFR<sub>TFS</sub> response across the four background noise listening conditions are displayed in Figure 26 below.



*Figure 26.* Subtracted grand average temporal waveforms for four listening conditions (Panel A: Clean condition, Panel B: +5 SNR condition, Panel C: 0 SNR condition, Panel D: -5 SNR conditions). Amplitude of the response is displayed along the y-axis and time is displayed along the x-axis

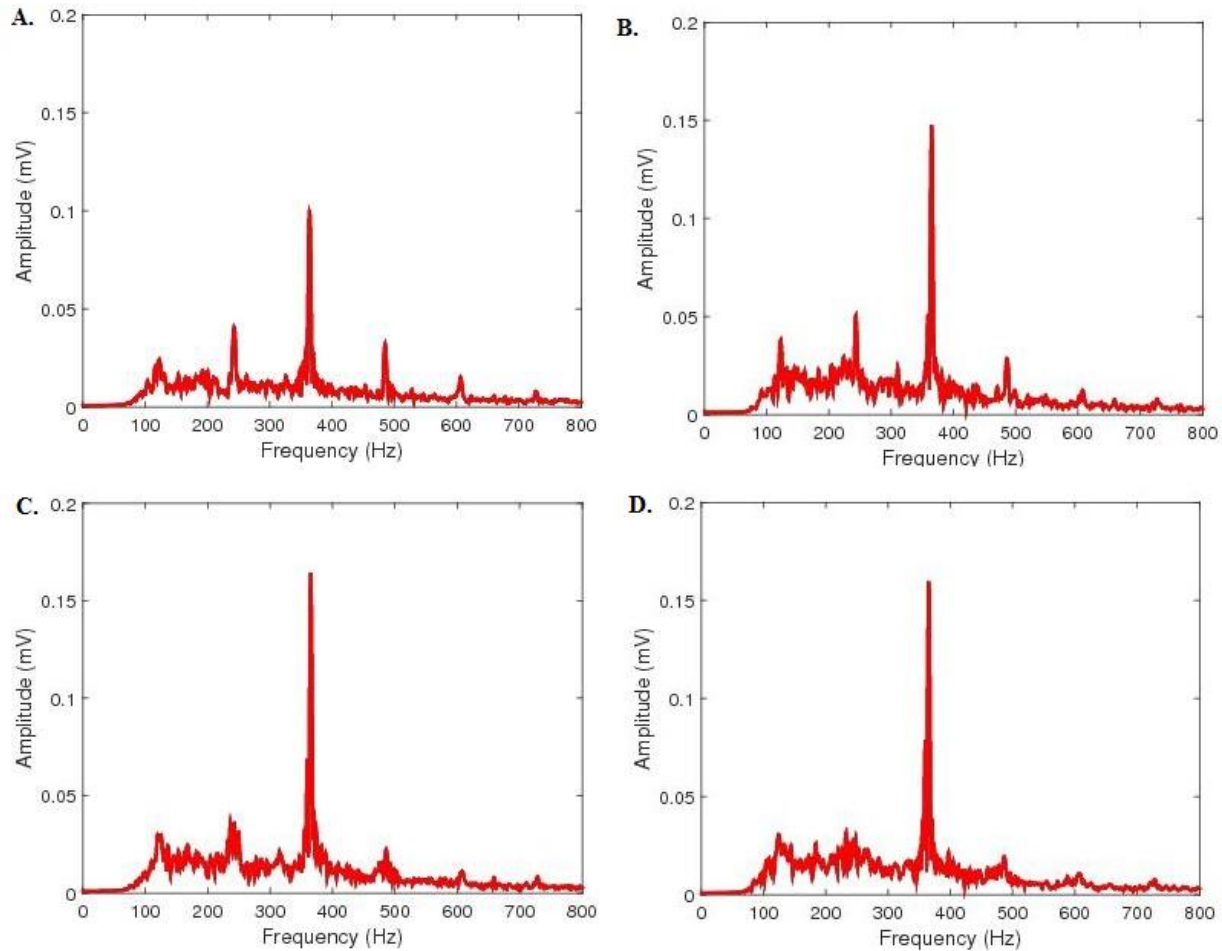
Visual inspection of the subtracted grand average temporal waveforms shows several interesting trends in mean amplitude and waveform morphology of the  $\text{FFR}_{\text{TFS}}$ . First, the waveform amplitude increases slightly as the severity of the background noise worsens. The overall amplitude in the clean condition is approximately .2 mV throughout the duration of the

response (250 ms). However, in the -5 dB SNR background noise listening condition, the amplitude reaches approximately .3 mV.

Second, overall waveform morphology seems to remain similar throughout all four listening conditions. These visual observations are not in agreement with the decrease in waveform amplitude and periodicity observed with the  $\text{FFR}_{\text{ENV}}$  waveforms as the severity of the background noise condition worsened.

### ***Spectral Analysis: Fast Fourier Transform (FFT)***

FFT analysis was used to convert  $\text{FFR}_{\text{TFS}}$  grand average temporal waveforms into the frequency domain for spectral analysis across all four listening conditions. The grand average FFT for each of these four background noise conditions are presented in Figure 27 below.



*Figure 27.* Subtracted grand mean FFT analysis for four listening conditions (Panel A: Clean condition, Panel B: +5 SNR condition, Panel C: 0 SNR condition, Panel D: -5 SNR conditions). Amplitude of the response is displayed along the y-axis and frequency is displayed on the x-axis.

Visual inspection of the grand average FFT for each background noise condition reveal several interesting trends in the spectral data as seen in Figure 27. First, the absolute magnitude of the first formant frequency (F1) at 360 Hz generally increases as the severity of the condition worsens: clean (.093 mV), + 5 dB SNR (.137 mV), 0 dB SNR (.149 mV), -5 dB SNR (.146 mV). The energy present at F1 is smallest in the clean condition. The F1 magnitude is essentially equal between the 0 SNR and -5 SNR listening condition (.149 mV and .146 mV, respectively).



Second, note that the peaks present at the other harmonic frequencies (whole number multiples of the F0) are much lower in energy than the energy present at F1 in all four listening conditions. For example, the energy present at the harmonic frequencies is .05 mV or less in all the four background noise conditions. The energy at these harmonics is substantially reduced in the -5 SNR background noise condition and essentially disappears in this most severe background noise listening condition.

### ***Quantitative Spectral Analysis: Individual data and Mean values***

The FFR<sub>TFS</sub> was initially analyzed in a similar manner as were the summed grand average temporal waveforms or F0 data. This analysis of the F1 data or the subtracted grand average temporal waveforms revealed an unexpected finding which was an increase in the mean F1 amplitude values as the severity of the condition worsened. For instance, the mean F1 amplitude values were: .093 mV, .137 mV, .149 mV, and .146 mV in the clean, +5 SNR, 0 SNR, and -5 SNR conditions, respectively.

It was hypothesized that the level of the noise floor may have been different in the four background noise conditions, which could have contributed to this unexpected F1 amplitude finding. The level of the noise floor is defined as the combination of both the resting electrical activity of the participant and the background noise that was presented simultaneously with the /u/ stimulus. The background noise is 5 dB greater than the signal in the -5 SNR condition and therefore more noise is present in the - 5 SNR condition as compared to the +5 SNR where it is 5 dB less than the signal. In order to explore this possibility, the level of the noise floor present in each of the four conditions was therefore measured. It was determined that these noise floor levels were in fact different across the test conditions. Specifically, the noise floor is smallest in

the clean condition (.015 mV), whereas the largest level of noise is present in the -5 SNR condition (.023 mV). The noise floors for the clean (.015 mV) and the +5 SNR condition (.019 mV) are less than the noise floors present in the 0 SNR (.023 mV) and -5 SNR condition (.023 mV).

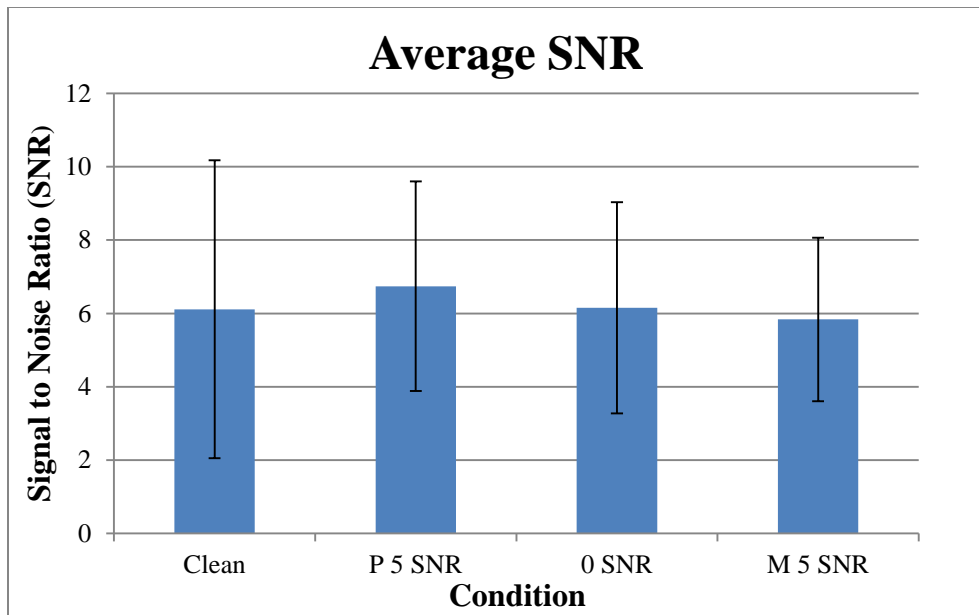
Due to these unexpected differences in the noise floor levels across test conditions, SNR values were calculated/derived for each test condition in order to compare the amplitude of the participants' subcortical response elicited by the /u/ stimulus to the overall noise floor. This technique in essence normalizes the effect of the noise floor across conditions. Note that these derived SNR calculations are different than the initial SNRs used to define how much background noise was simultaneously presented with the stimulus in the four different listening conditions. The table below summarizes the derived SNR values for each individual subject, as well as their mean and standard deviation values present in each of the four background noise conditions.

Table 6

*Individual Derived SNR measurements at the F1 and their Mean and Standard Deviation Values.*

	Condition				
	Clean	+ 5 SNR	0 SNR	-5 SNR	Difference
Subject 1	1.394	2.054	1.628	1.735	-.341
Subject 2	5.425	4.928	5.035	5.040	.385
Subject 3	7.743	6.674	8.782	6.781	.962
Subject 4	5.364	8.031	5.663	6.208	-.844
Subject 5	13.189	9.710	9.730	7.451	5.738
Subject 6	3.563	9.051	6.069	7.796	-4.233
<b>Mean</b>	<b>6.113</b>	<b>6.741</b>	<b>6.151</b>	<b>5.835</b>	<b>.278</b>
<b>SD</b>	<b>4.063</b>	<b>2.866</b>	<b>2.887</b>	<b>2.233</b>	<b>3.324</b>

As seen in Table 6 above, the largest mean derived SNR is present in the +5 SNR condition (6.11 dB) and the smallest mean SNR is present in the -5 SNR condition (5.83 dB). The variance was calculated for each background noise condition and is reflected in the standard deviation values. In general, the variability was fairly similar across test conditions, with the exception of the clean test condition. This finding supports the overall trend that the mean derived SNRs decrease as the severity of the listening condition worsens. The mean SNR values and standard deviations are visually depicted for the four test conditions in Figure 28 below.



*Figure 28.* Mean SNR measured for the four listening conditions (Clean, +5 SNR, 0 SNR, - 5 SNR) for the subtracted grand average temporal waveforms.

Individual participants' derived SNR values follow the same trend as the mean derived SNR values as seen in Table 6 above. Four out of the six participants had larger derived SNR values in the + 5 dB SNR listening condition than the derived SNR values in the -5 dB SNR condition. Subject one had derived SNRs which ranged from 1 to 2 dB across all four conditions. These similar SNRs and substantially smaller SNRs across test conditions for subject one are likely due to an error during data acquisition. Possible explanations for the trends seen in the mean SNR values will be addressed in the discussion section.

In summary, these results discussed above can be separated into two primary sections: the results of the  $FFR_{ENV}$  and results of the  $FFR_{TFS}$ . The primary trends seen in each are summarized below.

***FFR<sub>ENV</sub> results:***

1. As the severity of the background noise increases, the overall amplitude and periodicity of the grand average temporal waveform for the FFR<sub>ENV</sub> decreases.
2. Spectral analysis revealed a decrease in F0 magnitude as background noise increases, with the most substantial decrease occurring in the severe background noise condition.
3. The F0 magnitude is 50 % larger than the energy present at the harmonic frequencies for clean through moderate background noise conditions.
4. In contrast, for the severe condition, the energy present at the F0 and harmonics is approximately equal.

***FFR<sub>TFS</sub> results:***

1. The overall amplitude of the grand average temporal waveform increases as background noise conditions worsen.
2. Spectral analysis revealed an increase in F1 amplitude as the level of the background noise increases.
3. The energy present at the harmonics is much less than the energy present at the F1 in all test conditions; however, the harmonic amplitudes are smallest in the most severe listening condition.
4. Derived SNR values showed a decrease in SNR for F1 from the clean to -5 dB SNR background noise condition.

Collectively, the results for the FFR<sub>ENV</sub> and FFR<sub>TFS</sub> demonstrate a degradation of subcortical neural encoding of the F0 and F1 of the stimulus, respectively.

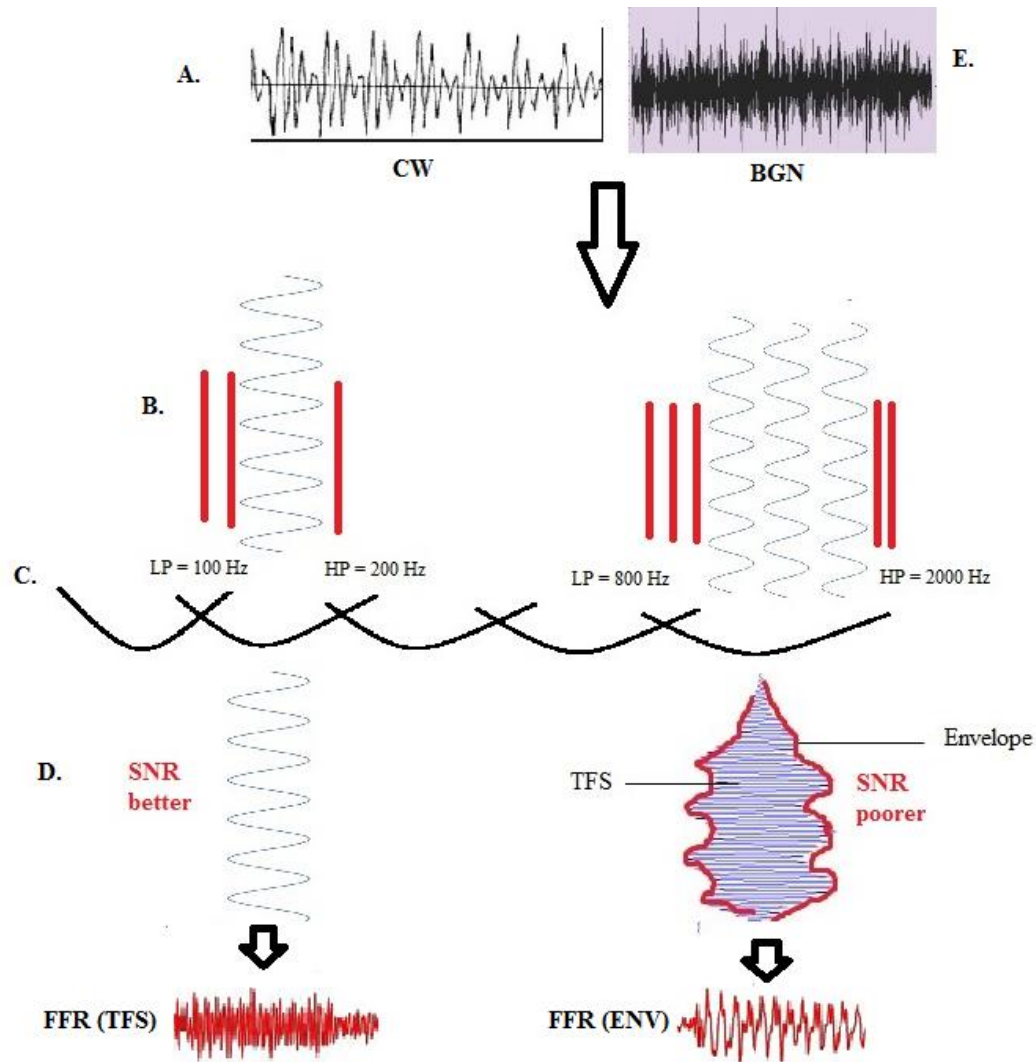
## **CHAPTER 5: DISCUSSION**

The present study sought to investigate the effects of different levels of background noise on subcortical neural encoding of the vowel stimulus /u/ in normal hearing adults. It was hypothesized that subcortical neural encoding of F0 and F1 would decrease as the severity of the background noise worsened. As expected, there was an overall degradation of FFR<sub>ENV</sub> as well as FFR<sub>TFS</sub> as the severity of background noise condition worsened. The most substantial degradation of FFR<sub>ENV</sub> and FFR<sub>TFS</sub> occurred in the -5 dB SNR background noise condition. A review of neural encoding at the level of the basilar membrane is required in order to understand why brainstem encoding of speech degrades in the presence of competing background noise.

### ***Overview of Neural Encoding at the Basilar Membrane***

As previously discussed in the literature review (p.6), a complex waveform (composed of an F0 and harmonics), such as the /u/ stimulus used in the present study, passes through overlapping band-pass filters along the basilar membrane. These band-pass filters separate the complex waveform into its different spectral components before the auditory system further processes the signal at structures higher in the auditory system (brainstem). The band-pass filters have a low frequency cut-off value and a high frequency cut-off value, which only allow a specific frequency or range of frequencies within the complex waveform to pass through the filter. Band-pass filters are arranged tonotopically, which means that the band-pass filters that pass low frequencies are located at the apical end of the cochlea, whereas the band-pass filters which pass high frequencies are located at the basal end of the cochlea. These band-pass filters are narrower at the apical end of the cochlea and wider at the basal end of the cochlea. Figure 29 below depicts a complex waveform (Step A). This complex waveform is passing through the

band-pass filters and is separated into its harmonic components (Step B). The narrow, low frequency band-pass filters only allow one frequency component to pass through. In contrast, the wide band-pass filters allow multiple high frequency components of the stimulus (such as higher harmonics) to pass through. Figure 29 below depicts this neural encoding at the basilar membrane (Step C) and the resulting output once the signals are divided into their respective “bands” (Step D). The lower frequency band- pass filters are narrower and therefore the output at these filters results in a resolved harmonic (simple pure tone). In contrast, higher frequency harmonics pass through a wide band-pass filter (such as a band-pass filter ranging from 800 Hz to 2000 Hz), resulting in unresolved harmonics. These unresolved harmonics take the shape of a complex waveform consisting of a slow varying envelope of  $F_0$  superimposed on a rapidly oscillating temporal fine structure (unresolved harmonic).



*Figure 29.* Depicts the tonotopic organization of band-pass filters along the basilar membrane and neural encoding of a complex waveform at the basilar membrane (Steps A through D). Background noise is also pictured entering the basilar membrane (bold red lines). Relative SNRs and corresponding FFR temporal waveforms at the output are displayed at each filter (Step E). (TFS= temporal fine structure, LP = Low pass frequency/cut off values, HP= high pass frequency/cut off values).



### *Effects of Background Noise on $FFR_{ENV}$*

When a complex speech signal is presented in background noise, the background noise is also processed through these cochlear band-pass filters at the same time that the target speech signal passes through the filters. Figure 29 above shows background noise (Step E) passing through the band-pass filters (as represented by the bolded red lines). Hence, the components of the speech signal must compete with background noise at the input and output of every auditory filter.

As described earlier, for the speech signal, the low frequency narrow filters yield only the resolved harmonics while the high frequency, wide filters yield unresolved harmonics. On the other hand, the level of background noise that enters and exits through these band-pass filters depends on the width of the filter. Narrow band-pass filters limit to an extent the amount of background noise that passes through, while the wide band-pass filters permit more background noise to pass through. As a result, in the case of the low frequency filter output, the level of the resolved harmonics belonging to the target signal are relatively stronger than the level of the background noise, resulting in a relatively strong SNR or spectral contrast. In contrast, at the output of the higher frequency filters, the level of the unresolved harmonics belonging to the target signal are not as strong as compared to the level of background noise, resulting in a weaker SNR or spectral contrast.

In addition to the difference between SNR values at the output of low frequency and high frequency band-pass filters, neural encoding of the F0 is not place specific. In other words, the majority of neural encoding for a low frequency component (F0) of the complex waveform does not take place at the output of the low-frequency band-pass filters. Instead, F0 encoding relies on

the neural phase-locking to the envelope of the unresolved harmonics at the output of the wide high frequency band-pass filters (Greenberg, Marsh, Brown & Smith, 1987). Note that this is visually depicted in Figure 29 above with the grand average temporal waveform for  $\text{FFR}_{\text{ENV}}$  below the high frequency band-pass filters. Smith, Marsh, Greenberg, and Brown (1978) demonstrated this phenomenon when they recorded the FFR on adults using a pure tone and a complex tone. Smith and colleagues presented a masking noise simultaneously with both the simple and complex tones that masked out the fundamental frequency of the stimulus. When the masking noise was presented simultaneously with the pure tone, the  $\text{FFR}_{\text{ENV}}$  was severely reduced. In contrast, when the masking noise was presented simultaneously with the complex tone, the  $\text{FFR}_{\text{ENV}}$  remained robust. The researchers proposed that  $\text{FFR}_{\text{ENV}}$  is governed by a “place” other than the one that approximates to the fundamental frequency of the stimulus, since the  $\text{FFR}_{\text{ENV}}$  remained robust for the complex tone with the missing fundamental (as created by the masking).

This “place” theory has also been suggested in a recent unpublished study conducted by Ananthakrishnan (2013). The FFR was recorded on normal hearing individuals and hearing impaired individuals with a high frequency sensorineural hearing loss using an /u/ stimulus. F0 encoding was compared between these two groups and was found to be more robust in normal hearing individuals than in individuals with a hearing impairment. In order to simulate the high frequency sensorineural hearing loss, a degraded /u/ stimulus, with high frequency information removed, was used to elicit the FFR in normal hearing participants. Interestingly, F0 encoding in normal hearing individuals to the degraded /u/ stimulus was similar to F0 encoding in hearing impaired individuals using the intact /u/ stimulus (Ananthakrishnan, 2013). This further indicates

that F0 encoding is dependent on high frequency information and without this information the F0 encoding is severely degraded.

Once these frequency components are encoded at the level of the basilar membrane, the auditory signal (resolved and unresolved harmonics) is then encoded in the auditory brainstem where neural phase locking occurs. The results from Smith et al. (1978) and Ananthakrishnan (2013) support the notion that neurons primarily phase lock to the envelope at the output of the wide, high frequency band-pass filters (Sayles & Winter, 2008). Therefore the  $\text{FFR}_{\text{ENV}}$  is a result of neural phase-locking to the slow varying envelope superimposed on a rapidly oscillating TFS and corresponds to the F0 of the stimulus.

Collectively, given that: 1. The SNR increases at the wider band-pass filters when a signal is presented to a listener in the presence of background noise and; 2. The  $\text{FFR}_{\text{ENV}}$  is mediated by neural encoding of the envelope at these broad (high frequency) band-pass filters, a degradation in  $\text{FFR}_{\text{ENV}}$  corresponding to the F0 of the stimulus should be expected as function of background noise condition.

The results for the current study do in fact indicate that neural encoding of the F0 as represented by the  $\text{FFR}_{\text{ENV}}$  did degrade with increasing levels of background noise. This effect is likely due to the underlying physiological model of neural encoding described above. The findings from the present study are consistent with previous literature. Li and Jeng (2011) found an overall decrease in F0 encoding, or F0 amplitude, in normal hearing adults with varying levels of background noise starting at +12 dB SNR and decreasing to a -12 dB SNR in 6 dB increments. F0 amplitudes significantly decreased from the most favorable condition (+12 dB SNR) to the least favorable background noise condition (-12 dB SNR) (Li & Jeng, 2011).

Additionally, Song et al. (2010) stated that F0 amplitude significantly decreases for all subjects at a + 10 dB SNR using different types of background noise.

The second most notable finding in the present study was that  $\text{FFR}_{\text{ENV}}$  did not substantially decrease until the worst background noise condition (-5 dB SNR). These results are consistent with Li and Jeng (2011) findings. These researchers reported that F0 encoding was tolerant to noise and remained robust in the no noise condition, the +12 dB SNR condition, and the + 6 dB SNR condition. It was not until 0 dB SNR, that Li and Jeng reported that the F0 amplitude substantially decreased. The results from the present study were also consistent with Russo et al. (2004) and Song et al. (2010) who found that F0 amplitude remained robust at positive SNR values.

In the current study, the energy at the F0 is substantially larger in relative magnitude to the harmonics, for the positive SNR conditions. It is equal in magnitude to the energy at the harmonics in the -5 dB SNR background noise condition. Results from animal studies are consistent with the present study. For example, Cunningham et al. (2002) recorded near-field neural responses on animals to a speech syllabus in varying levels of background noise at the inferior colliculus (a known generator of the FFR). These investigators reported that neural firing rates were most affected at 0 dB SNR. Collectively, the results of these animal and human studies are consistent with the present study, which indicate that the  $\text{FFR}_{\text{ENV}}$  is least affected when the signal is greater in magnitude than the noise.

### ***Effects of Background Noise on $\text{FFR}_{\text{TFS}}$***

The effect of BGN on  $\text{FFR}_{\text{TFS}}$  can also be explained by the functioning of the cochlear filter bank and subsequent neural encoding of speech sounds and background noise, as described

above. Leek and Summers (1996) proposed that background noise leads to “broadly tuned cochlear filtering” even in normal hearing adults. This may cause low frequency filters, that are supposed to be narrow, to widen, resulting in the production of *more* unresolved harmonics than would occur in quiet listening conditions. Hence, this broadly tuned cochlear filtering results in an increase in background noise at the output of even the low frequency filters, with no subsequent increase in the energy at the desired signal. This in turn would lead to further reductions in the spectral contrasts, even at low frequency harmonics that are typically resolved.

This finding that spectral contrast in background noise is reduced at low frequency filters due to a broadening of the of the filter is important because several studies have proposed that the  $\text{FFR}_{\text{TFS}}$  represents neural encoding of resolved harmonics at the output of these low frequency filters. In the same unpublished study that provided evidence that  $\text{FFR}_{\text{ENV}}$  is mediated by high frequency information at the output of broad band-pass filters, there were also findings that supported that  $\text{FFR}_{\text{TFS}}$  is mediated by harmonics passed through these narrow, low frequency band-pass filters (Ananthakrishnan, 2013). The FFR was recorded on normal hearing and hearing impaired individuals with a high frequency sensorineural hearing loss using an intact /u/ stimulus. F1 encoding was found to be more robust for normal hearing listeners than for hearing impaired listeners. The high frequency information was then removed from the /u/ stimulus and presented to the normal hearing listeners. F1 encoding of the degraded /u/ stimulus for normal hearing listeners was still greater than F1 encoding of the intact /u/ stimulus for the hearing impaired individuals. Thus, these results indicated that high frequency regions and the information at these regions do not contribute to F1 encoding. Note that this is visually depicted in Figure 29 above with the grand average temporal waveform for  $\text{FFR}_{\text{TFS}}$  (below the low

frequency band-pass filters). These results are in contrast to  $\text{FFR}_{\text{ENV}}$ , which does depend on high frequency information for robust neural encoding.

Overall, the results from the present study showed a decrease in SNR for F1 as the background noise condition worsens. The smallest derived SNR was present in the most severe background noise condition (-5 dB SNR) and can be likely attributed to larger amounts of noise passing through these low frequency band-pass filters and therefore decreasing the internal SNR (Leek & Summers, 1996). These results are consistent with results found in the literature. Russo et al. (2004) reported that the amplitude of F1 decreased in the presence of background noise. These investigators also commented that F1 magnitude was significantly above the noise floor in the quiet and +5 dB SNR background noise condition, with the higher harmonics remaining below the noise floor. In this study, the energy at F1 was also larger than the energy at the harmonics for the clean, 0 and +5 dB SNR conditions. The only exception to this was for the severe background noise condition (-5 dB SNR). In this condition, the higher harmonics were substantially lower in amplitude and likely submerged in the noise floor.

There is limited research that examines the effects of background noise on F1 neural encoding. Much of the literature pertaining to the effects of background noise on F1 or other formant frequencies consists of studies that use behavioral results rather than FFR recordings. For example, Hedrick and Nabelek (2004) used an FFT on several different vowel temporal waveforms to view F2 magnitude in the frequency domain. Results from that study indicate that a relative decrease in F2 magnitude in background noise caused vowel misperception. One could speculate that poor performance on behavioral testing, such as speech perception in noise, could be a result of an underlying decrease in SNR or reduced spectral contrast between the F2 magnitude and noise floor.

### ***Clinical Implications***

The majority of the published literature uses animal models to investigate the effects of background noise on neural encoding of an auditory signal. To date, much of the literature using human participants focuses on behavioral results such as speech perception in noise. For example, the literature reports that decreased magnitude at F1 as a function of increasing background noise levels can lead to incorrect vowel identification (Nabalek, 1988). The current study which employed the FFR to investigate the effects of background noise on neural encoding may serve as a link between these animal and human studies to help explain underlying neurological mechanisms that may contribute to poor speech perception in noise. Furthermore, since hearing in noise is a common complaint for those with a hearing impairment, results from this study can be used to help investigate the effects of background noise on neural encoding in hearing impaired listeners.

### ***Limitations***

The largest limitation of the current study was the number of participants. A limited sample size of six subjects only allowed for descriptive measures to be reported in the present study. A larger sample size of approximately 12 would have allowed the analysis to go beyond just stating the trends in the data. Statistical analysis on data using a larger sample size may have yielded a significant finding between the various background noise conditions. Another limitation of the study was that a single vowel stimulus was used. The /u/ stimulus is a limited representation of speech. Whereas neural responses to words or sentences presented in background noise may provide more insight regarding the neural encoding of speech that is occurring at subcortical regions of the brain in normal, everyday noisy listening environments.

### ***Further Studies***

Future studies may wish to compare subjects' performance on both behavioral tasks as well as their electrophysiologic results obtained using the FFR. This type of comparison was used previously by Anderson et al. (2010) who recorded the FFR using a /da/ stimulus in quiet and with a + 10 dB SNR. The subjects' FFR results were then compared to the subjects' performance on the speech in noise tests. An increase in FFR latency was related to poor performance on the Hearing in Noise Test (HINT) (Anderson et al., 2010). Future studies analyzing the FFR together with behavioral studies could shed light on whether a degradation of FFR<sub>ENV</sub> and FFR<sub>TFS</sub> contributes to poor speech intelligibility in noise. Furthermore, future studies could compare the effects of different types of background noise on subcortical neural encoding using the FFR. Steady state noise has been shown to impact speech intelligibility more negatively than fluctuating noise such as speech shaped noise (Festen & Plomp, 1990). Lastly, future research could investigate the effects of age and hearing impairment on the FFR.



## APPENDIX A

### Informed Consent Form

I, \_\_\_\_\_, agree to participate in a study entitled “The Effects of Background Noise on Subcortical Neural Encoding of Speech Stimuli in Normal Hearing Adults,” which is being conducted by Audiology doctoral student Laura Somers of the Audiology, Speech-Language Pathology, and Deaf Studies Department at Towson University. The purpose of study is to evaluate auditory neural encoding ability in the presence of background noise. It is hoped that the information obtained from this study may provide insight into why normal hearing individuals experience listening difficulty in noisy listening environments.

As a participant, I understand that I am committing to approximately two, 3.0 hour test sessions. During these sessions, I will be asked to relax and sit comfortably on a recliner while electrophysiological recordings are taken from my scalp using scalp electrodes.

I have been informed that any information obtained in this study will be recorded with a unique code number that will allow the primary investigators to determine my identity. If any information from this study is used for publication or presentation at a professional meeting, I will in no way be identified and my name will not be used.

I understand that there is no personal risk or discomfort directly involved with this research as standard electrophysiological techniques will be used. I understand that my participation is voluntary, and that I am free to withdraw my consent and discontinue participation in this study at any time. If I decide to withdraw from this study it will in no way affect any future services I may obtain from the Department of Audiology, Speech Language Pathology and Deaf Studies.

If I have any questions or problems that arise in connection with my participation in this study, I should contact Dr. Saradha Ananthakrishnan, the thesis chair by email (sananthakrishnan@towson.edu) or at 410-704-5903 or Dr. Deb Garland the IRB Chairperson at (410) 704-2236.

---

(Date)

---

(Signature of Participant)

---

(Date)

---

(Investigator)

THIS PROJECT HAS BEEN REVIEWED BY THE INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN PARTICIPANTS AT TOWSON UNIVERSITY.

## APPENDIX B



**APPROVAL NUMBER: 15-A054**

To: Randi Cropper  
2 Waterway Ct APT 2D  
Towson MD 21286

From: Institutional Review Board for the Protection of Human  
Subjects Scot McNary, Member

Date: Monday, February 23, 2015

RE: Application for Approval of Research Involving the Use of  
Human Participants

Office of Sponsored Programs  
& Research

Towson University  
8000 York Road  
Towson, MD 21252-0001

t. 410 704-2236  
f. 410 704-4494

Thank you for submitting an Application for Approval of Research Involving the Use of Human Participants to the Institutional Review Board for the Protection of Human Participants (IRB) at Towson University. The IRB hereby approves your proposal titled:

*The difference in the effects of adverse listening conditions on subcortical neural encoding of speech stimuli in normal hearing adults*

If you should encounter any new risks, reactions, or injuries while conducting your research, please notify the IRB. Should your research extend beyond one year in duration, or should there be substantive changes in your research protocol, you will need to submit another application for approval at that time.

We wish you every success in your research project. If you have any questions, please call me at (410) 704-2236.

CC: S. Ananthkrishnan; P. Korezak  
File

## References

- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing research*, 245(1), 35-47.
- Akhoun, I., Gallégo, S., Moulin, A., Ménard, M., Veuillet, E., Berger-Vachon, C., ... & Thai-Van, H. (2008). The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme/ba/in normal-hearing adults. *Clinical Neurophysiology*, 119(4), 922-933.
- Ananthakrishnan, Saradha, "Neural Encoding of Complex Signals in the Healthy and Impaired Auditory Systems" (2013). Open Access Dissertations. Paper 198.
- Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural timing is linked to speech perception in noise. *The Journal of Neuroscience*, 30(14), 4922-4926.
- Assmann, P., & Summerfield, Q. (2004). The perception of speech under adverse conditions. In *Speech processing in the auditory system* (pp. 231-308). Springer New York
- Atcherson, S. R., & Stoody, T. M. (Eds.). (2012). *Auditory electrophysiology: a clinical guide*. Thieme.
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral cortex*, 19(11), 2699-2707.
- Berglund, B., Hassmen, P., & Job, R. S. (1996). Sources and effects of low-frequency noise. *The Journal of the Acoustical Society of America*, 99(5), 2985-3002.
- Biacabe, B., Chevallier, J. M., Avan, P., & Bonfils, P. (2001). Functional anatomy of auditory brainstem nuclei: application to the anatomical basis of brainstem auditory evoked potentials. *Auris Nasus Larynx*, 28(1), 85-94.

- Borden, G. J., Harris, K. S., & Rafael, L. J. (1984). *Speech science primer* Williams and Wilkins. Baltimore, MD, 1.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, 47(2), 236-246.
- Chiappa, K. H. (Ed.). (1997). *Evoked potentials in clinical medicine*. Lippincott Williams & Wilkins.
- Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. *Hearing research*, 264(1), 48-55.
- Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clinical Neurophysiology*, 112(5), 758-767.
- Davis, H., & Hirsh, S. K. (1976). The audiometric utility of brain stem responses to low-frequency sounds. *International Journal of Audiology*, 15(3), 181-195.
- Drullman, R. (1995). Speech intelligibility in noise: relative contribution of speech elements above and below the noise level. *The Journal of the Acoustical Society of America*, 98(3), 1796-1798.
- Du, Y., Kong, L., Wang, Q., Wu, X., & Li, L. (2011). Auditory frequency-following response: A neurophysiological measure for studying the “cocktail-party problem”. *Neuroscience & Biobehavioral Reviews*, 35(10), 2046-2057.
- Emanuel, D. C., & Letowski, T. (2009). *Hearing science*. Wolters Kluwer Health/Lippincott Williams and Wilkins.

- French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *The journal of the Acoustical society of America*, 19(1), 90-119.
- Galbraith, G. C. (1994). Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 92(4), 321-330.
- Galbraith, G. C., & Doan, B. Q. (1995). Brainstem frequency-following and behavioral responses during selective attention to pure tone and missing fundamental stimuli. *International journal of psychophysiology*, 19(3), 203-214.
- Galbraith, G. C., Threadgill, M. R., Hemsley, J., Salour, K., Songdej, N., Ton, J., & Cheung, L. (2000). Putative measure of peripheral and brainstem frequency-following in humans. *Neuroscience letters*, 292(2), 123-127.
- Gelfand, S. (2001). *Essentials of audiology* (2nd ed.). New York: Thieme
- Gelfand, S. A. (Ed.). (2009). *Essentials of audiology*. Thieme.
- Gong, Q., Xu, Q., & Sun, W. (2013). Design and implementation of frequency-following response recording system. *International journal of audiology*, 52(12), 824-831.
- Greenberg, S., Marsh, J. T., Brown, W. S., & Smith, J. C. (1987). Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hearing Research*, 25(2), 91-114.
- Hall, J. W. (2007). *New handbook of auditory evoked responses* (pp. 123-136). Pearson.
- Hairston, W. D., Letowski, T. R., & McDowell, K. (2013). Task-related suppression of the brainstem frequency following response. *PLoS One*, 8(2), e55215.

- Hedrick, M. S., & Nabelek, A. K. (2004). Effect of F2 intensity on identity of /u/ in degraded listening conditions. *Journal of Speech, Language, and Hearing Research*, 47(5), 1012-1021.
- Henry, K. S., & Heinz, M. G. (2012). Diminished temporal coding with sensorineural hearing loss emerges in background noise. *Nature neuroscience*, 15(10), 1362-1364.
- Jewett, D. L., & WILLISTON, J. S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain*, 94(4), 681-696.
- Katz, J. (Ed.). (1994). Handbook of clinical audiology.
- Kraus, N., & Nicol, T. (2005). Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. *Trends in neurosciences*, 28(4), 176-181.
- Krishnan, A. (1999). Human frequency-following responses to two-tone approximations of steady-state vowels. *Audiology and Neurotology*, 4(2), 95-103.
- Krishnan, A. (2002). Human frequency-following responses: representation of steady-state synthetic vowels. *Hearing research*, 166(1), 192-201.
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25(1), 161-168.
- Krishnan, A., Bidelman, G. M., Smalt, C. J., Ananthakrishnan, S., & Gandour, J. T. (2012). Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*, 50(12), 2849-2859.
- Krizman, J., Skoe, E., & Kraus, N. (2010). Stimulus rate and subcortical auditory processing of speech. *Audiology and Neurotology*, 15(5), 332-342.

- Kvale, M. N., & Schreiner, C. E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. *Journal of Neurophysiology*, 91(2), 604-612.
- Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *The Journal of Neuroscience*, 29(18), 5832-5840.
- Leek, M. R., Dorman, M. F., & Summerfield, Q. (1987). Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 81(1), 148-154.
- Leek, M. R., & Summers, V. (1996). Reduced frequency selectivity and the preservation of spectral contrast in noise. *The Journal of the Acoustical Society of America*, 100(3), 1796-1806.
- Lehmann, A., & Schönwiesner, M. (2014). Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. *PloS one*, 9(1), e85442.
- Li, X., & Jeng, F. C. (2011). Noise tolerance in human frequency-following responses to voice pitch. *The Journal of the Acoustical Society of America*, 129(1), EL21-EL26.
- Marsh, J. T., Worden, F. G., & Smith, J. C. (1970). Auditory frequency-following response: neural or artifact?. *Science*, 169(3951), 1222-1223.
- Marsh, J. T., Brown, W. S., & Smith, J. C. (1974). Differential brainstem pathways for the conduction of auditory frequency-following responses. *Electroencephalography and clinical neurophysiology*, 36, 415-424.
- Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27(2), 338-352.

Moller, A. R. (2006). *Hearing: anatomy, physiology, and disorders of the auditory system*.

Academic press.

Moore, B. C. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology*, 9(4), 399-406.

Nábělek, a K. (1988). Identification of vowels in quiet, noise, and reverberation: relationships with age and hearing loss. *The Journal of the Acoustical Society of America*, 84(2), 476–84.

Nábělek, a K., & Dagenais, P. a. (1986). Vowel errors in noise and in reverberation by hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 80(3), 741–8.

Moushegian, G., Rupert, A. L., & Stillman, R. D. (1973). Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalography and clinical neurophysiology*, 35(6), 665-667.

Picton, T. W. (2010). *Human auditory evoked potentials*. Plural Pub Incorporated.

Pickett, J. M. (1957). Perception of Vowels Heard in Noises of Various Spectra. *Journal Of The Acoustical Society Of America*, 29(5), 613. doi:10.1121/1.1908983

Rinne, T., Balk, M. H., Koistinen, S., Autti, T., Alho, K., & Sams, M. (2008). Auditory selective attention modulates activation of human inferior colliculus. *Journal of neurophysiology*, 100(6), 3323-3327.



- Robinson, A. J. (2008). *Clinical electrophysiology: electrotherapy and electrophysiologic testing*. Lippincott Williams & Wilkins.
- Roeser, R. J., Valente, M., & Hosford-Dunn, H. (Eds.). (2007). *Audiology: diagnosis* (Vol. 1). Thieme.
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, 115(9), 2021-2030.
- Sinha, S. K., & Basavaraj, V. (2010). Speech evoked auditory brainstem responses: a new tool to study brainstem encoding of speech sounds. *Indian Journal of Otolaryngology and Head & Neck Surgery*, 62(4), 395-399.
- Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and hearing*, 31(3), 302.
- Smalt, C. J., Krishnan, A., Bidelman, G. M., Ananthakrishnan, S., & Gandour, J. T. (2012). Distortion products and their influence on representation of pitch-relevant information in the human brainstem for unresolved harmonic complex tones. *Hearing research*, 292(1), 26-34.
- Smith, J. C., Marsh, J. T., & Brown, W. S. (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and clinical neurophysiology*, 39(5), 465-472.
- Smith, J. C., Marsh, J. T., Greenberg, S., & Brown, W. S. (1978). Human auditory frequency-following responses to a missing fundamental. *Science*, 201(4356), 639-641.
- Sohmer, H. H., Pratt, H. H., & Kinarti, R. R. (1977). Sources of frequency following responses (FFR) in man. *Electroencephalography And Clinical Neurophysiology*, 42(5), 656-664.

- Song, J. H., Banai, K., Russo, N. M., & Kraus, N. (2006). On the relationship between speech- and nonspeech-evoked auditory brainstem responses. *Audiology and Neurotology*, 11(4), 233-241.
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Perception of speech in noise: neural correlates. *Journal of Cognitive Neuroscience*, 23(9), 2268-2279.
- Swaminathan, J., Krishnan, A., Gandour, J. T., & Xu, Y. (2008). Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *Biomedical Engineering, IEEE Transactions on*, 55(1), 281-287.
- Wever, E. G., & Bray, C. W. (1930). The nature of acoustic response: The relation between sound frequency and frequency of impulses in the auditory nerve. *Journal of experimental psychology*, 13(5), 373.
- Young, E. D., & Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. *The Journal of the Acoustical Society of America*, 66(5), 1381-1403.

## CURRICULUM VITAE

**NAME:** Laura Somers

**PROGRAM OF STUDY:** Audiology

**DEGREE AND DATE TO BE CONFERRED:** Doctor of Audiology (Au.D.), 5/2016

**SECONDARY EDUCATION:** Mt. Hebron High School

Collegiate institutions attended	Dates	Degree	Date of Degree
St. Vincent College— <i>Latrobe, PA</i>	8/2008 - 5/2012	B.A.	5/2011
Towson University— <i>Towson, MD</i>	8/2012 - 5/2016	Au.D.	Anticipated 5/2016

**PROFESSIONAL POSITIONS HELD:**

**\*\*All were clinical internships in accordance with Towson University's Audiology Doctoral Program**

A and A Hearing Group  
1813 York Road  
Lutherville-Timonium, MD 21093  
Doctoral Student Clinician, January 2015—May 2015

ENTAA Care  
203 Hospital Drive, Suite 200  
Glen Burnie, MD 21061  
Doctoral Student Clinician, June 2014 — December 2014

Howard County Diagnostic Center  
5451 Beaverkill Road  
Columbia, MD 21044  
Doctoral Student Clinician, January 2014 – May 2014

Towson University Speech, Language and Hearing Center  
One Olympic Place  
Towson, MD 21204  
Doctoral Student Clinician, January 2013—December 2013

