

Running head: FREQUENCY FOLLOWING RESPONSE: EXAMINING
VARIABILITY

TOWSON UNIVERSITY
OFFICE OF GRADUATE STUDIES

FREQUENCY FOLLOWING RESPONSE (FFR): EXAMINING VARIABILITY
OF THIS COMPLEX RESPONSE

by

Lindsay R. Roberts

A thesis

presented to the faculty of

Towson University

in partial fulfillment

of the requirement for the degree

Doctor of Audiology

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

Towson University
Towson, MD 21252

April 2017

FREQUENCY FOLLOWING RESPONSE: EXAMINING VARIABILITY

TOWSON UNIVERSITY
OFFICE OF GRADUATE STUDIES
THESIS APPROVAL PAGE

This is to certify that the thesis prepared by Lindsay Roberts entitled Frequency Following Response: Examining the Variability of this Complex Response has been approved by the thesis committee as satisfactorily completing the thesis requirements for the degree Doctor of Audiology.

A. Saradh 4.17.2017
Chairperson, Thesis Committee Signature Saradha Ananthakrishnan Date

Peggy Korczak 4/17/17
Committee Member Signature Peggy Korczak Date

Elise Smith 4/17/17
Committee Member Signature Elise Smith Date

Janet V Delany 4-20-17
Dean of Graduate Studies Janet Delany Date

Acknowledgements

First, I would like to thank my thesis committee chair, Dr. Saradha Ananthakrishnan for sharing her expertise in the frequency-following response and providing guidance and patience throughout the process of writing this doctoral thesis. I would also like to thank my committee members Dr. Peggy Korczak and Dr. Elise Smith for their encouragement, suggestions, and support. Special thanks to Laura Grinstead, my co-investigator, for providing accountability and answers to my many questions throughout this process. Last of all, I would like to thank my friends and family who helped contribute to this project, as a participant or as encouragement during the writing process.

FREQUENCY FOLLOWING RESPONSE: EXAMINING VARIABILITY

Abstract

Frequency Following Response (FFR): Examining Variability of this Complex Response

Lindsay R. Roberts

One aspect of the frequency-following response (FFR) which reflects neural representation of pitch at the level of the brainstem that has proven difficult to explain is the significant amount of variability seen in the response, even within a group of normal hearing listeners (Ruggles, Bharadwaj, & Shinn-Cunningham, 2011). Such response variability has been attributed to a number of factors including but not limited to experience or training induced plasticity, gender and age of the subject, “hidden hearing loss”, stimulus parameters, analysis parameters, etc. The specific aims of this study were twofold: 1) to examine the effect of stimulus frequency on the FFR to low frequency pure tone stimuli (250, 500 and 1000 Hz) in a group of normal-hearing individuals, and 2) to examine subject-related variability in the brainstem neural encoding strength, specifically, if and how gender and varying degrees of musical experience are reflected in the neural encoding of a 500 Hz pure tone stimuli. The FFR was recorded to 250, 500, and 1000 Hz pure tone stimuli, in 30 normal hearing 19-29-year-old young adults (males = 12). A total of 4000 sweeps were collected at each frequency per participant, and data

FREQUENCY FOLLOWING RESPONSE: EXAMINING VARIABILITY

were analyzed by gender, and self-reported number of years of formal musical training. Results revealed a significant effect of stimulus frequency on FFR spectral amplitude. Additionally, there was a significant effect of gender, with females displaying greater spectral amplitudes than males at 500 Hz. The relationship between years of musical training and FFR amplitude was non-significant. These findings suggest that neural representation of pitch is greater for low frequency pure tone stimuli, when compared to frequencies, such as 1000 Hz, that are closer to the upper limits of phase locking in the human auditory brainstem. Additionally, the differences between males and females observed in the FFR amplitudes at 500 Hz, suggest that gender differences may be present in sub-cortical neural representations of simple auditory stimuli.

Table of Contents

List of Figures *viii*

CHAPTER 1: LITERATURE REVIEW *1*

 Overview of Auditory Evoked Potentials 1

 Basic Acoustics Review 5

 Neural Encoding 16

 History and Neural Generators of the FFR 22

 Subject Parameters 25

 Subject state. 25

 Subject age 26

 Gender. 27

 Auditory experience. 28

 Stimulus Parameters 31

 Stimulus intensity. 32

 Stimulus rate. 33

 Stimulus polarity. 34

 Recording Parameters 35

 Electrode montage. 35

 Sampling rate. 36

 Analysis window. 36

 Sweeps. 37

 What Does the FFR Tell Us? 37

 Research Question 44

CHAPTER 2: METHODOLOGY *47*

 Participants 47

 Materials and Measures 47

 Stimuli 47

 Procedure. 47

 Data analysis. 48

 Statistical analysis. 49

CHAPTER 3: RESULTS *50*

FREQUENCY FOLLOWING RESPONSE: EXAMINING VARIABILITY

CHAPTER 4: DISCUSSION..... 56

 Effects of Frequency on FFR Amplitude 56

 Effects of Gender on FFR Amplitude 56

 Effects of Musical Training on FFR Amplitude 59

 Clinical Implications..... 62

 Study Limitations and Future Directions..... 62

APPENDICES.....64

 APPENDIX A: Informed Consent Form 64

 APPENDIX B: Case History Form..... 65

 APPENDIX C: Goldsmith Musician Sophistication Index (Page 5)..... 66

 APPENDIX D: IRB Approval 67

References 68

CURRICULIM VIATE 78

FREQUENCY FOLLOWING RESPONSE: EXAMINING VARIABILITY

List of Figures

<i>Figure 1.</i> This image shows the simple waveforms of three pure tones, that are added together to form a complex waveform, labeled D.....	6
<i>Figure 2.</i> This image is an example of a complex aperiodic waveform.....	8
<i>Figure 3.</i> The images on the left display the simple (A) and complex (B) waveforms in the time domain, the right are graphed in the frequency domain ...	11
<i>Figure 4.</i> This figure displays the difference between steady state and time varying stimuli.....	13
<i>Figure 5.</i> This image is a spectrogram of the CV speech syllable /da/.....	15
<i>Figure 6.</i> Phase locking behavior for the auditory neurons in response to a 1000 Hz stimulus.	18
<i>Figure 7.</i> Basilar membrane filtering of a pure tone.....	19
<i>Figure 8.</i> Basilar membrane filtering of a complex harmonic tone.....	21
<i>Figure 9.</i> Original FFR recording to a steady-state English back vowel /u/ with an $F_0=120$ Hz and an $F_1=360$ Hz.....	39
<i>Figure 10.</i> The black waveform is the Envelope FFR, resulting from the addition of the responses to the condensation and rarefaction polarity stimuli.	40
<i>Figure 11.</i> FFT of the FFR_{ENV} (black temporal waveform in Figure 10). This FFT reveals a large magnitude at 120 Hz, the F_0 of the vowel /u/.	42
<i>Figure 12.</i> This FFT reveals a primary spectral component with a large magnitude at approximately 365 Hz, the first formant frequency of the vowel /u/.....	43
<i>Figure 13.</i> Grand Averaged FFR_{TFS} waveforms for the three stimulus frequencies tested plotted as a function of time (ms).....	51
<i>Figure 14.</i> This figure shows the grand averaged FFR_{TFS} FFT magnitude measured in μV , as a function of frequency in Hz.	52
<i>Figure 15.</i> This graph shows the grand averaged FFR_{TFS} FFT magnitude measured in μV , plotted as a function of pure tone stimulus frequency in Hz.	53
<i>Figure 16.</i> This graph illustrates the average FFR_{TFS} FFT magnitude at 500 Hz measured in μV , plotted as a function of gender and number of years of musical training.	55

CHAPTER 1: LITERATURE REVIEW

Overview of Auditory Evoked Potentials

An evoked potential arises with the generation of a compound action potential following the stimulation of sensory receptors past their resting threshold (Passmore, Murphy, & Lee, 2014). Evoked potentials are able to reflect the functional integrity of both sensory and motor systems (Kiylioglu, Parlaz, Akyildiz, & Tataroglu, 2015). Specifically, evoked potentials are a measure of the electrical signals elicited by the brain recorded in response to stimulation of specific sensory nerve pathways including the visual system, peripheral nerve, and auditory system. The visual evoked potential (VEP) is an electrical signal produced by the brain in response to a time-locked visual stimulus; VEP can be used to detect eye diseases in patients, including but not limited to: glaucoma, diabetic retinopathy, macular degeneration, color blindness, and the loss of peripheral vision (Paulraj, Subramaniam, Yaccob, Adom, & Hema, 2015). The somatosensory evoked potential (SEP), measures the electrical activity response at the surface of the skin following the controlled stimulation of the peripheral nerve; SEP can detect problems with the patients' spinal cord, as well as numbness and weakness of the extremities (Passmore et al., 2014). The auditory evoked potential (AEP) is the electrical signal elicited from the brain in response to time-locked auditory stimuli (Paulraj et al., 2015).

AEPs are the changes in the electrical activity of the brain in response to acoustical stimuli (Picton, 2010; Plourde, 2006). AEPs contain both positive and negative waves (deflections) that follow the auditory stimulus in a time-locked manner (Plourde, 2006). Evoked potential equipment use bioamplifiers to make the recorded

response signal large enough for adequate signal processing due to the small amplitude of the unamplified response (Burkard & Don, 2015). The amplified electrical activity must then be converted into a binary format (via analog to digital conversion) to be processed, filtered, and averaged by the digital computer (Burkard & Don, 2015). The series of waves produced by the AEP reflect the auditory ability of the individual, which allow AEPs to be used clinically for hearing screenings, intraoperative monitoring, and to estimate hearing thresholds in difficult-to-test populations (Burkard & Don, 2015; Paulraj et al., 2015). AEPs are also used in research, and the frequency following response (FFR) is one such AEP that is being studied, making its way to clinical applications. A thorough review of AEPs is required before further discussion of the FFR.

Auditory evoked potentials are classified by their latency, recording location, source, time, and stimulus properties. The latency of an AEP is defined as the time between the onset of the stimulus and the onset of the response (Picton, 2011). Specifically, AEPs are classified as fast (1-15ms), middle (10-50ms), slow (30-500ms) or late (200-1000ms) based on the latency of the response after the stimulus presentation (Picton, 2011). Fast AEP responses include waves I-VII of the auditory brainstem response (ABR) and the FFR. The middle latency response (MLR) is comprised of three waves (Na, Pa, and Nb), and as it is clearly defined in its name, is a middle latency response occurring from 10-50 ms post stimulus onset. The slow AEP responses include the cortical sustained potential and the mismatch negativity (MMN) and these responses comprise a portion of the late AEP responses (Paulraj et al., 2015; Picton, 2011).

Evoked potentials can be recorded from either a near-field or far-field location. When a potential is recorded in close proximity to the source of the bioelectric activity, it

is considered a near-field recording (Stegeman, Roeleveld, Dumitru, & King, 1997). Near-field recordings with needle electrodes placed within the brainstem are often completed in animal studies with recordings occurring throughout the brain (Jewett & Willison, 1971). Far-field recordings are more frequently seen in research and take place when the electrodes are placed at a distance from the generator of the response, such as on the scalp. Typically, AEP responses measured in humans are a far-field response recorded from the scalp using non-invasive surface electrodes. The electrodes and electrode paste serve as the interface between the scalp and the electrical instrumentation (Burkard & Don, 2015).

Following the presentation of an acoustic stimulus to the auditory system, electrical activity can be recorded in the cochlea, auditory nerve, auditory brainstem, thalamus, and auditory cortex (Burkard & Don, 2015). The source of the response has a direct effect on the latency of the response; responses from the lower structures (more caudal) will have shorter onset latencies, while responses from the structures that are higher up in the auditory system (more rostral) will have longer latencies (Burkard & Don, 2015). For example, waves IV-VI of the ABR reflect the bioelectrical activity from the superior olivary complex to the lateral lemniscus, whereas, waves Na-Pb-Nb of the MLR arise from the thalamus and primary auditory cortex (Fowler & Horn, 2012).

AEPs are also classified by the timing and duration of the stimulus and response. Transient responses are evoked by brief, non-sustained stimulus features and follow a change in the stimulus (the onset or offset of the stimulus) and occur when the stimulus presentation is slow enough that the response wears off before the next stimulus is presented (Picton, 2011; Plourde, 2006; Skoe & Kraus, 2010). Transient responses are

characterized by fast response peaks and include responses such as the ABR, MLR and MMN (Picton, 2011; Skoe & Kraus, 2010). The sustained response is evoked continually throughout the presentation of the stimulus and originates from the periodic features phase-locked to the period of the stimuli (Picton, 2011; Skoe & Kraus, 2011). The FFR is an example of a sustained brainstem response that follows the frequency of the input signal. Steady-state AEPs, such as the auditory steady-state response, are evoked when the stimulus is changing regularly, and occur when the stimulus rate is fast enough to create an overlap of the individual transient responses (Picton, 2011; Plourde, 2006).

The AEP comprises both the exogenous and endogenous components of the brain. Early latency responses like the ABR are considered exogenous responses because they are sensitive to the acoustics characteristics and intensity of the incoming signal, while nearly insensitive to the subject's attention to the stimuli (Tremblay & Clinard, 2015). Endogenous responses provide information on higher cognitive processes of the brain and are affected by the attention of the patient to the stimulus (Picton, 2011; Munivrana & Mildner, 2013). The mismatch negativity response, related to change detection, is generally considered endogenous because the subject must be attentive to the changes occurring in the properties of the stimuli (Stapells, 2009; Picton, 2011).

Based on the above classification schemes, the frequency-following response (FFR) is a fast, scalp-recorded (i.e. far field), sustained, exogenous brainstem response that reflects phase-locked activity of brainstem neurons. A better understanding of basic acoustics, speech science and neural encoding is required before discussing the FFR in further detail.

Basic Acoustics Review

Acoustics, a branch of physics, is defined as the study of sound. Sound is graphically represented by waveforms, which are functions that represent changes of a physical quantity over of time (Emanuel & Letowski, 2009). There are two types of acoustic waveforms, simple and complex. Simple waveforms, including pure tones, have only one frequency component, while complex waveforms are made up of multiple frequency components. A single completion of a waveform pattern is called a cycle. The frequency of a simple periodic tone is defined as the number of cycles that occur in one second. A complex waveform is created when at least two pure tones of different frequencies are added together; this process is called waveform synthesis and can be seen in the bottom of Figure 1 (Emanuel & Letowski, 2009). Figure 1 shows the waveforms of 3 different pure tones (labeled waveforms A-C, respectively); when combined they form the complex waveform displayed at the bottom of the figure, labeled D.

Complex waveforms can be further broken down into aperiodic and periodic waveforms. Aperiodic complex waveforms, as seen in Figure 2, are random and have no observable repeatable pattern. There are two main types of complex aperiodic waveforms: noise and transients. Noise is a random sequence of events that is made up of an enormous number of unrelated components (Emanuel & Letowski, 2009). A transient is a burst of noise that lasts only for a short period of time (e.g. slamming a book on a table).

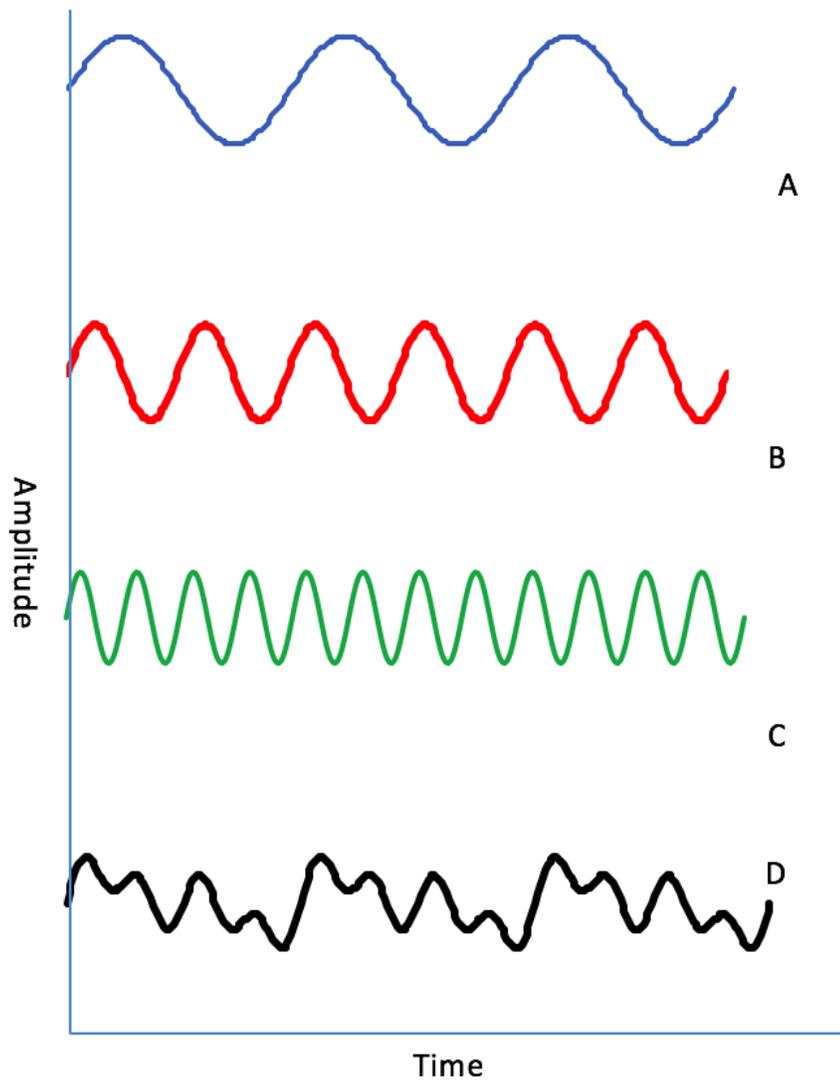


Figure 1. The above image shows the simple waveforms of a three different pure tones, labeled A, B, and C, that are added together to form a complex waveform, labeled D.

A periodic complex waveform (as seen in the bottom image in Figure 1) is one where the pattern of vibration, a complete cycle, however simple or complex, repeats itself in the time domain (Borden, Harris, & Raphael, 1994). The greatest common factor of all the frequencies involved in a complex waveform is known as the fundamental frequency. For example, in a complex waveform made up of 100, 200, 400 and 800 Hz tones, the fundamental frequency would be 100 Hz. In this example, all of the components of the complex waveform are whole-number multiples of the fundamental frequency, and are called harmonics (Emanuel & Letowski, 2009). Harmonics are generally denoted by F_n , where n indicated the integer multiple. For instance, in the example described above, the complex stimulus contains the harmonics F_1 , F_2 , F_4 , and F_8 . In some cases, there is a missing fundamental frequency; this occurs when the greatest common factor of a complex waveform is not equal to a frequency component present in the waveform itself. For instance, a missing fundamental of 500 Hz would be present in a complex waveform made up of 1000, 1500 and 2000 Hz tones.

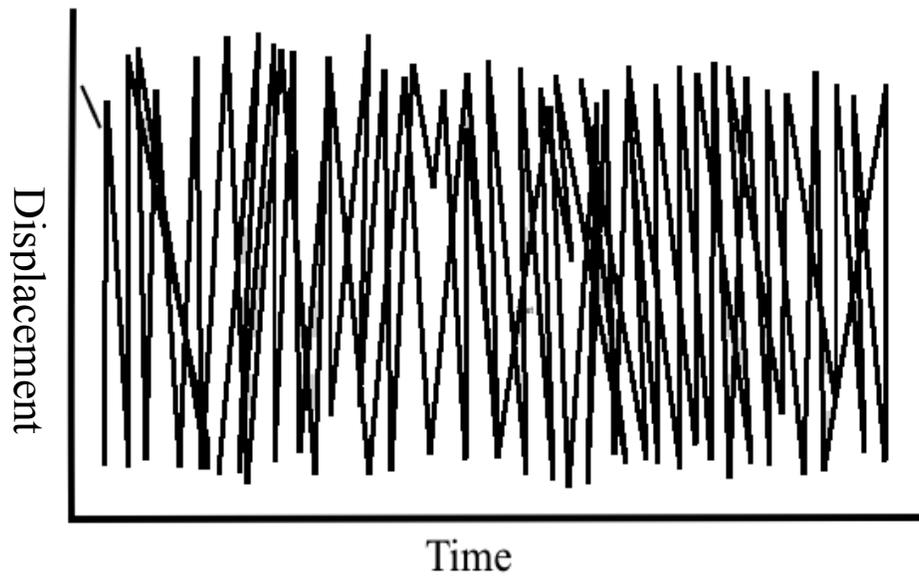


Figure 2. This image is an example of a complex aperiodic waveform. Note that there is no observable or repeatable pattern.

Just as simple waveforms can be combined together to form complex waveforms, by a process called waveform synthesis, complex waveforms can be broken down into their individual frequency components through waveform analysis (Emanuel & Letowski, 2009). As it was previously defined, waveforms are functions that represent the changes of a physical quantity in the time domain. Figures 1 and 2 display the behaviors of the vibrations across time in the time domain, displaying amplitude as a function of time. By simply looking at the graphical display of a complex waveform it is generally impossible to determine the frequency components. In order to display the frequency components of a waveform, Fast Fourier Transform (FFT) analysis is used to show the waveform energy (amplitude) present at individual frequency components making up the complex waveform. (Emanuel & Letowski, 2009).

FFT analysis is one of the waveform analyses that are frequently used to extract the frequency information from a complex signal recorded in the time domain (Clark & Ohlemiller, 2008). Although the specific mathematics involved in the FFT are beyond the scope of this paper, it should be known that FFT is able to transform the value of a function and convert it to another related function (Clark & Ohlemiller, 2008). When FFT is used to break down a complex waveform, the spectral components reveal the fundamental frequency or F_0 of the waveform and any harmonics of the F_0 , if they are present in the waveform. The vertical lines displayed in the spectrum are the spectral components labeled C and D in Figure 3. The height of the vertical line on the x-axis of the spectra indicates the amount of energy present at that frequency, which is graphed along the x-axis. A simple sine tone displayed in the frequency domain would only have a single vertical line as it only has one spectral component (see top image in Figure 3)

(Emanuel & Letowski, 2009). In contrast, a complex waveform with three individual frequency components would have three vertical lines to display energy contained at the three different frequency components (labeled B in Figure 3) (Emanuel & Letowski, 2009). Note that the x-axes for the waveforms on the left in Figure 3 are time and the x-axes for the spectrum on the right are frequency.

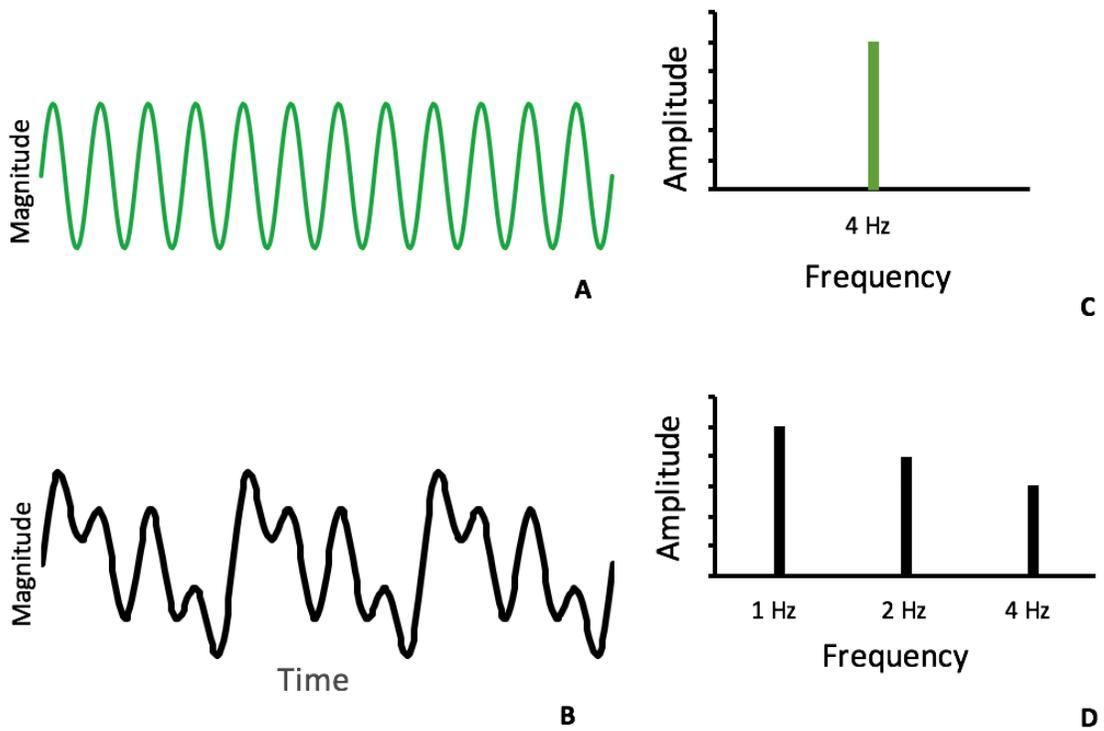


Figure 3. The images on the left display the simple (A) and complex (B) waveforms in the time domain, the images on the right illustrate the result of the FFT conducted on the time waveforms. The vertical lines represent the individual spectral components, contained in time waveforms, as a function of frequency. This figure was adapted from Emanuel & Letowski, 2009.

Speech is a complex periodic stimulus. In humans, our voice serves as the vehicle for speech to occur and its acoustic features are represented by the fundamental frequency (F_0) and the formant patterns (Nan, Skoe, Nicol & Kraus, 2015). Speech sounds are produced when air coming from the lungs causes the vocal folds to vibrate (Skoe & Kraus, 2010). The vibrations of the vocal folds in the larynx are the source of speech production and the rate of this vocal fold vibration determines the F_0 . The sound emitted by the larynx contains energy at the F_0 as well as integer multiples of the F_0 also known as harmonics (Belin, 2006). This harmonically-rich sound is filtered (modified) by the speech articulators, cavities and tissues above the larynx such as the teeth, tongue and lips, to form different speech sounds (Belin, 2006; Skoe & Kraus, 2010). Different configurations of the speech articulators alter the resonance properties of the vocal tract and cause some harmonics of the F_0 to be amplified and others to be attenuated. This filtering creates resonance frequencies, called formants, that reinforce the energy at certain frequencies and are observed as peaks in the speech spectrum (Belin, 2006; Skoe & Kraus, 2010).

A speech signal differs greatly from a simple periodic pure tone. One of the differences is that pure tones and other simple periodic stimuli are steady state stimuli meaning their frequency does not change over time. In contrast, real life speech sounds such as diphthongs, and consonant vowel syllables (/da/) are time varying with changes in frequency occurring over time (Skoe & Kraus, 2010). As seen in Figure 4, when time has elapsed, the frequency of the steady state stimuli stays the same, while, the frequency of the time varying stimuli changes from 250 to 500 Hz. The FFR can be recorded to both steady state and time varying stimuli.

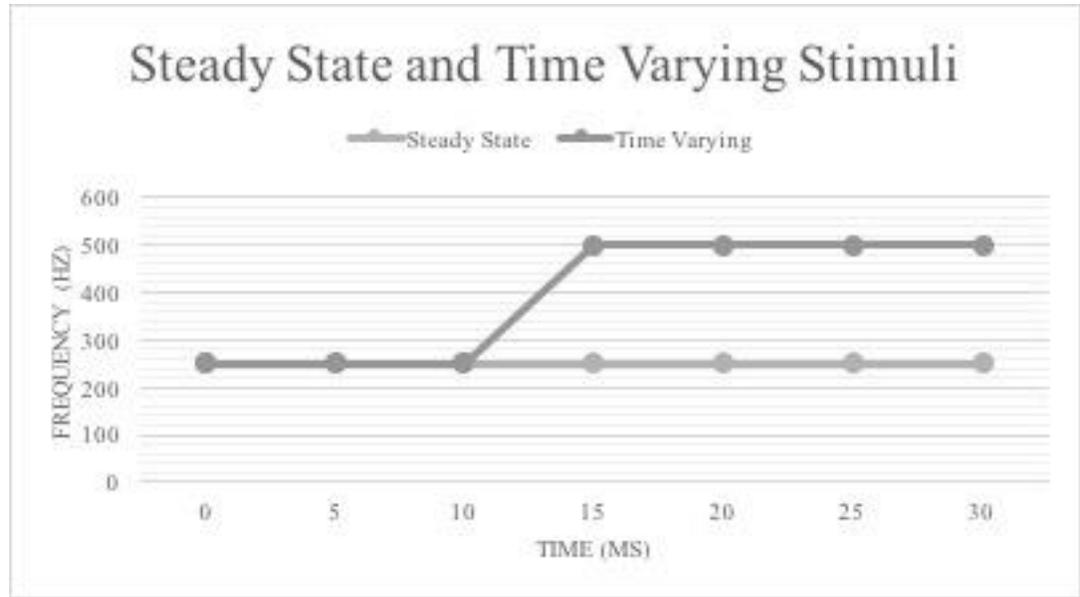


Figure 4. This figure displays the difference between steady state and time varying stimuli. Note that as time changes the frequency of the steady state stimuli stays the same while the frequency of the time varying stimuli changes from 250 to 500 Hz.

A spectrum as seen in the right panel of Figure 3, is beneficial for viewing the frequency components of a complex signal at one point in time. In order to visually represent the frequency components of a complex speech waveform over time, a spectrogram is used. Spectrograms are three dimensional graphs of the fundamental, harmonic and formant frequencies of a stimulus as a function of time (Skoe & Kraus, 2010). As seen in Figure 5, the x-axis of the spectrogram is time and the frequency is displayed on the y-axis. The intensity of the frequency components in the speech sound /da/ are depicted by the darkness of the frequencies shown. The formants of the speech sound are the darkest frequency bands towards the bottom of the figure and are labeled T1 and T2.

Now that we have reviewed the components of a spectrum in the frequency domain, we can begin to discuss the possible filters that can be applied to these spectra. A filter is a mechanism that adjusts the spectrum of its input and passes only a section of the spectrum to the output (Clark & Ohlemiller, 2008). There are four main types of acoustic filters: low pass, high pass, band pass and band reject. Low-pass filters will pass all frequencies below the pre-determined cut off frequency. Similarly, the high-pass filter will pass all frequencies above the cut off frequency. Band-pass filters will pass only the frequencies between two different cut off frequencies, rejecting the frequencies above or below the “band”. Band-reject filters are the opposite of a band pass filter, rejecting the frequencies between the two cut off frequencies and accepting all other frequencies.

The basic concepts of acoustics and speech science can be applied to the application of neural encoding that takes place at different points along the auditory system starting at the basilar membrane.

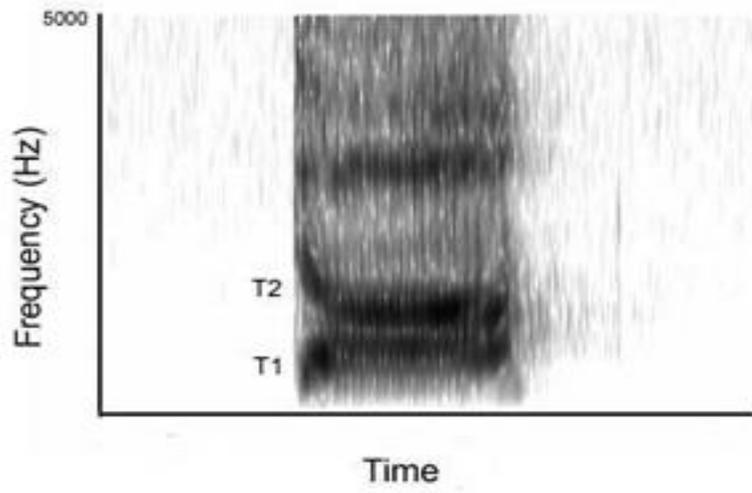


Figure 5. This image is a spectrogram of the CV speech syllable /da/. The formant frequencies are labeled T1 and T2.

Neural Encoding

The neural encoding of pitch has been described with two different theories throughout the literature: spectral and temporal, with some models integrating both theories (Walker, Bizley, King, & Schnupp, 2011). Spectral theories suggest that the frequency discrimination of pitch is represented by the place of maximal activation along the tonotopic organization of the auditory system (Walker et al., 2011). These spectral or “place” models rely on the organization of auditory filters that are present in the auditory periphery, with changes in the frequency represented by the displacement of a different portion of the system (Moore & Glasberg, 1989; Walker et al., 2011). Temporal theories, also described as rate models, state that the pitch of the signal is extracted from the timing of the neural impulses that are phase locked to the incoming waveform of the sound (Walker et al., 2011).

Neural phase locking occurs at all levels of the auditory system. Specifically, phase locking occurs when low frequency stimuli evoke neural responses with peaks occurring at intervals equal to the period of the stimulus (Moushegian, Rupert, & Stillman, 1973). For example, as seen in Figure 6, for a stimulus with a frequency of 1000 Hz, the period of the stimulus ($1000\text{ms}/1000\text{ Hz}$) is 1 ms. For this 1000 Hz stimulus, the neural discharges will occur in time with the positive deflecting peaks at 1 ms intervals (i.e. 1 ms, 2 ms, 3 ms, 4 ms, etc...). At the level of the basilar membrane, auditory nerve fibers phase lock to frequencies up to about 5000 Hz (Sek & Bukala, 2013; Walker et al., 2011). As we ascend the auditory pathway, from the basilar membrane to the auditory midbrain, thalamus and cortex, phase locking ability reduces, and the greatest frequency at which phase locking occurs decreases (Cariani, 1999). The

upper limits of phase locking in the inferior colliculus at the level of the midbrain has been found to be 2000 Hz (Moushegian et al., 1973; Rose, Brugge, Anderson, & Hind, 1967). In the primary auditory cortex, fundamental frequencies up to 400 Hz are recorded in the synchronized phase locked responses of the local auditory neurons (Cariani, 1999). The recording of the FFR requires synchronized neural activity in the form of phase locking (Aiken & Picton, 2008). The FFR primarily arises from generators in the inferior colliculus and thus reliably reflects the temporal discharge activity of the auditory neurons to the F_0 for frequencies up to 1500 Hz (Greenberg et al., 1987; Moushegian et al., 1973; Skoe & Kraus, 2010).

Temporal neural encoding of pure tones, speech and other complex periodic stimuli is a complex process. The basilar membrane can be thought of as a series of overlapping band pass filters. These filters follow the tonotopic organization of the cochlea with filters near the base of the cochlea tuned to high frequencies and those at the apex tuned to the low frequencies. Each filter is tuned to a specific frequency, also known as the characteristic frequency of that filter. As depicted in Figure 7, when a simple pure tone stimulus, labeled A, passes through the outer and middle ears and enters the cochlea, a single “place” along the basilar membrane will be maximally excited (Sek & Bukala, 2013). Because the pure tone is only made up of one frequency component, it will pass through an auditory filter tuned to that particular frequency (Figure 7B). Since the pure tone passed through the filter alone, the signal at the output of the filter is said to be *resolved*, also known as resolved temporal fine structure(TFS) (Figure 7C).

Phase Locking

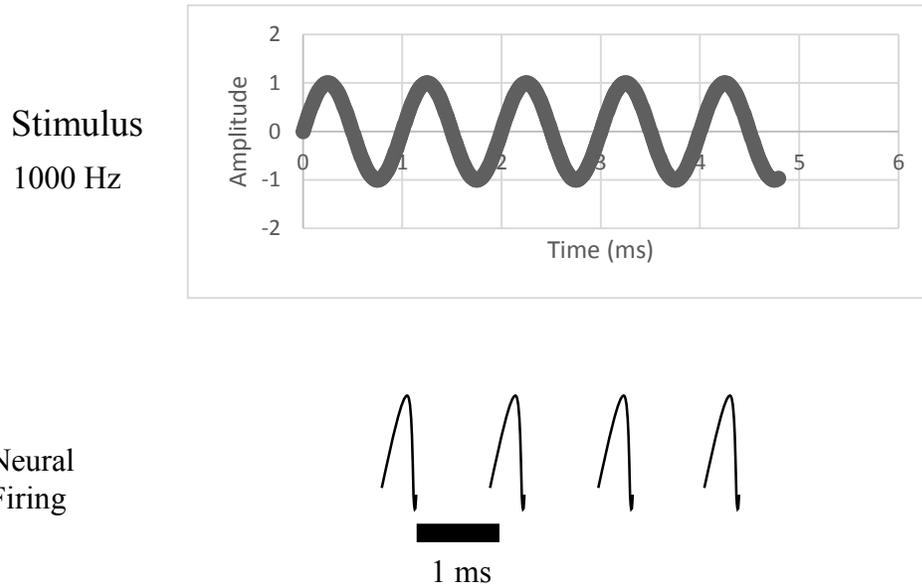


Figure 6. Phase locking behavior for the auditory neurons in response to a 1000 Hz stimulus. Note that the neural peaks are occurring at 1ms intervals which is equal to the period of the 1000 Hz stimulus.

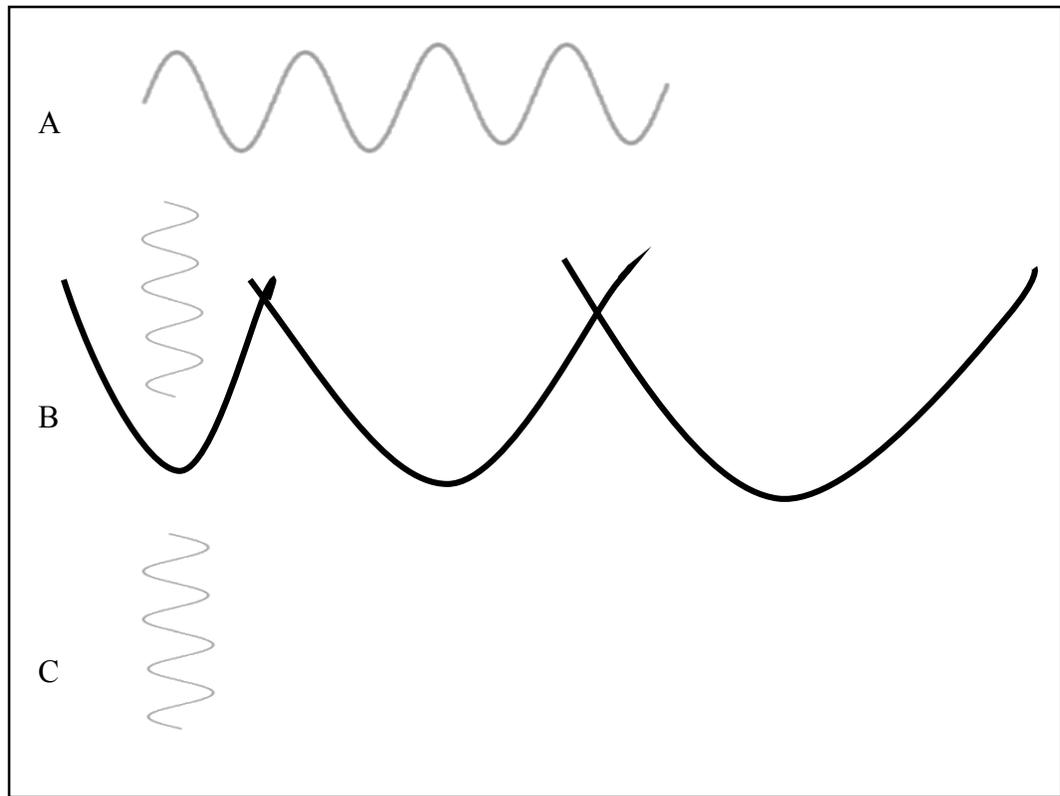


Figure 7. Basilar membrane filtering of a pure tone. A. pure tone. B. Tonotopically organized basilar membrane filters. As the sine wave passes through its own filter the output is resolved. C. resolved harmonic product of basilar membrane filtering.

This process is a little more complicated when complex signals, which are made of multiple frequencies, are considered. The spacing of the harmonic components of the signal is linear (the distance between two components is the same), however, the spacing of filters along the basilar membrane is roughly logarithmic (the width of an auditory filter increases as the center frequency of the filter increases) (Sayles & Winter, 2008; Sek & Bukala, 2013). Hypothetically, a band pass filter best tuned to 150 Hz may allow frequencies from 135-165 Hz to pass through the filter, where a band pass filter tuned to 4000 Hz may allow frequencies from 3000-5000 Hz to pass (Sayles & Winter, 2008; Sek & Bukala, 2013). When a harmonic complex tone (Figure 8A), composed of many sine waves, arrives at the cochlea, many places of maximal displacement will be observed along the basilar membrane. The low frequency harmonics of the complex tone will each pass through filters tuned to those particular frequencies because the spacing between them is larger than the width of the auditory filter (Figure 8B); these are the resolved harmonics (Figure 8C) (Sayles & Winter, 2008; Sek & Bukala, 2013). In case of high frequency harmonics, the distance between successive harmonics is smaller than the width of the auditory filter. Here, several different high frequency harmonics may pass together through a single broad filter tuned to a high frequency (Figure 8D); these are considered *unresolved* harmonics) (Sayles & Winter, 2008; Sek & Bukala, 2013). The unresolved harmonics that passed through the single filter form another complex waveform that contains a slowly varying envelope (Figure 8E, gray portion) superimposed on a rapidly varying TFS (Figure 8E, blue portion) (Moore, 2008; Sek & Bukala, 2013). In the case of complex stimuli, the neural envelope is modulated at a frequency corresponding to stimulus F_0 (Greenberg, Marsh, Brown, & Smith, 1987).

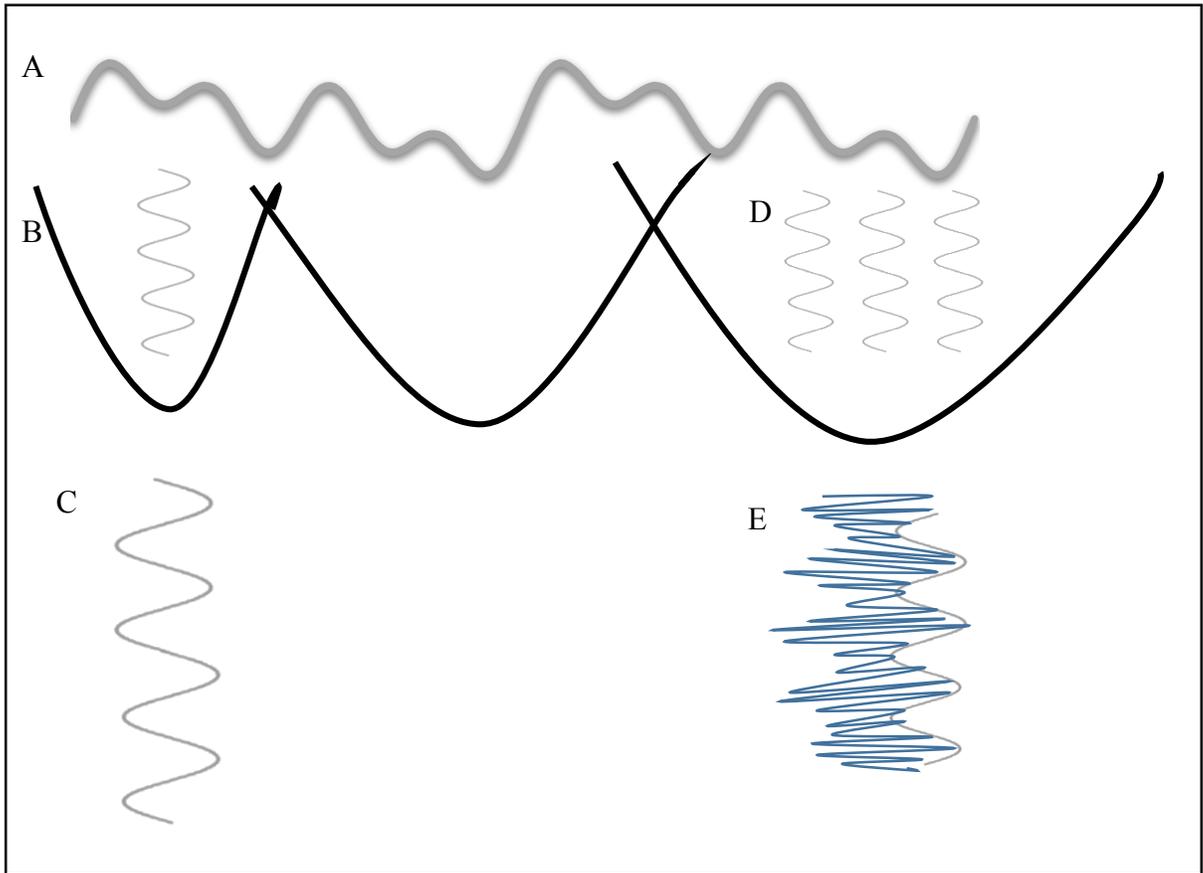


Figure 8. Basilar membrane filtering of a complex harmonic tone. A. Complex harmonic tone. B. Tonotopically organized basilar membrane filters. Low harmonics of the signal pass through their own filters. And the result is a resolved harmonic (C). D. Higher harmonics may pass through one single filter. E. The result of the filtering of higher harmonics is a complex waveform that contains a slowly varying envelope (E, gray portion) superimposed on a rapidly varying temporal fine structure (E, blue portion).

History and Neural Generators of the FFR

With a foundation of basic acoustics and the neural encoding of auditory stimuli, the FFR can be discussed more specifically. The frequency following response was accidentally discovered in 1930 by Wever and Bray, although their discovery was not confirmed by Worden and Marsh until 1968. Wever and Bray (1930) recorded nerve impulses directly from the auditory nerve in response to sound in anesthetized cats. These initial recordings from the auditory nerve produced action currents that were audible in the receiver as sounds that were perceptually identical to the original stimulus. This is now known as the cochlear microphonic. When the active electrode was placed into the tissues of the brainstem, the responses that were perceptually similar to the original auditory stimulus continued to be recorded. Further, when the cochlea was removed the recordings remained, suggesting that electrical processes accompany the conduction of nerve impulses when these action potentials recorded from the brainstem (Wever & Bray, 1930).

Worden and Marsh (1968) confirmed Wever and Bray's findings and were the first to describe this cochlear microphonic-like neural component with a F_0 equal to the frequency of the stimulus. Because this component followed the frequency of the stimulus it was coined the frequency following response (FFR). In their near-field recordings in 17 cats, Worden and Marsh, (1968), concluded that the FFR was indeed a true neural response and not simply an artifact. They reported that the FFR is recordable up to the level of the inferior colliculus.

Moushegian and colleagues, (1973), were the first to record the frequency following response with scalp electrodes in humans. In five normal hearing listeners,

Moushegian et al., (1973) recorded the FFR to low frequency sinusoid tones ranging from 250-2000 Hz. They found that these low frequency sounds elicited predictable neural waveforms recorded from the scalp with response peaks occurring at intervals equal to the period of the stimulus frequency. For example, 500 and 1000 Hz signals generated neural peaks that occurred at 2 and 1 ms intervals, respectively (Moushegian et al., 1973).

Initially, the FFR was confused with the cochlear microphonic, a potential generated by the cochlear hair cells. Although both the FFR and the cochlear microphonic mimic the temporal waveform of the incoming stimuli, there are noticeable differences between the two potentials (Chandrasekaran & Kraus, 2010). The cochlear microphonic is easily distinguished from the FFR in that the cochlear microphonic occurs almost simultaneously with the stimulus presentation while the FFR occurs from approximately 6-10 ms post stimulus onset, suggesting an origin rostral to the cochlea (Chandrasekaran & Kraus, 2010; Greenberg et al., 1987). Additionally, the FFR and the cochlear microphonic are differentially affected by increases in stimulus rate and intensity. Where the cochlear microphonic is largely unaffected by increases in stimulus rate, the FFR shows latency shifts with increasing rates (Skoe & Kraus, 2010; Worden & Marsh, 1968). The amplitude of the cochlear microphonic increases linearly with increases in stimulus intensity, while the amplitude of FFR reaches a plateau at supra-threshold intensity levels (Skoe & Kraus, 2010). It has been established by Worden and Marsh, (1968), among others, that the FFR is a true neural response and not simply the cochlear microphonic. However, due to their similarities, care must be taken to distinguish the cochlear microphonic from FFR recordings (Skoe & Kraus, 2010).

Although the cochlea has been ruled out as the origin of the FFR, the precise neural generator of the FFR has been a subject of debate in the FFR literature. In a series of two experiments, Smith, Marsh and Brown (1975) explored the brainstem origin of the scalp-recorded FFR. Their first experiment compared the onset latencies of the FFR recorded directly from the brainstem nuclei (near-field recordings) and the scalp (far-field recordings) in cats to those recorded from the scalp in humans. The recordings from the cochlear nucleus (CN), medial superior olive (MSO), and inferior colliculus (IC) in the cats revealed mean onset latencies of 2.0, 3.1 and 5.4 ms, respectively. The mean onset latency of the FFR recorded from the IC was found to closely approximate the mean latency of 5.8 ms recorded from the scalp in cats. When they compared the average latency of the scalp recorded FFR in the cat to that of their human subjects, they found no significant differences in latency (Smith et al., 1975). In their second experiment, Smith and colleagues (1975) measured the FFR before, during and after cooling of the IC from within the brainstem and the scalp. Measurements from the left MSO, left and right IC and scalp revealed that bilateral cooling of the IC greatly reduced or eliminated the FFR recordings from within the IC and from the scalp. After considering the findings of these two studies, Smith et al., (1975) concluded that at moderate stimulus intensities, the inferior colliculus was the primary source of the scalp recorded FFR.

In stark contrast to the conclusion of Smith and colleagues, (1975), Gardi, Merzenich and McKean, (1979), suggest that the inferior colliculus plays no major role in the generation of the scalp-recorded FFR. They found that when they completely aspirated the IC in 6 cats, that the amplitude of the scalp-recorded FFR was largely unaffected. In their analyses of the scalp-recorded FFR in cats through sectioning of

different brainstem nuclei, Gardi and colleagues suggested that the FFR has multiple generators (1979). Gardi et al., (1979) found that the cochlear nucleus provides approximately 50% of the amplitude in the scalp-recorded FFR in cats. Further, these researchers reported that FFR contributions from the cochlea and SOC account for approximately 25% and 20% of the response amplitude, respectively (Gardi et al., 1979).

Although contradictory in their findings regarding a specific neural origin, the cooling and ablation studies reviewed above both suggest a strong brainstem origin for the scalp-recorded FFR (Chandrasekaran & Kraus, 2010). Further supporting this conclusion, the ABR has a latency of about 1.5-5.5 ms, whereas, the FFR latency is approximately 6 ms, suggesting a more rostral brainstem origin. It is possible that the FFR reflects multiple neural generators located throughout the rostral brainstem, including the CN, IC, and MSO (Chandrasekaran & Kraus, 2010).

Subject Parameters

In general, aspects of scalp recorded auditory evoked potentials can be affected by the subject's attention or sleep, age, gender and auditory experience.

Subject state. Galbraith and Doan (1995) observed the effects of attention on FFR amplitudes in a group of 32 normal hearing listeners in response to a 400 Hz pure tone and a complex tone with a missing fundamental frequency. Subjects were asked to attend to the stimuli in one ear, while ignoring the stimuli in the other ear. Additionally, they performed a behavioral discrimination task where they were asked to press a button when they perceived the change in stimuli (varying duration or intensity) in the left or right "attend" ear. They observed larger FFR amplitudes in the channel that required the participant's attention compared to the channel that they were ignoring (Galbraith &

Doan, 1995). These results suggest that selective attention can modify the human auditory evoked brainstem response (Galbraith & Doan, 1995). These findings are reinforced by Lehmann and Schönwiesner, (2014), who concluded that selective attention to a single vowel sound modulates the human FFR based on combined spatial and frequency cues as well as frequency cues alone. Due to the observed changes in FFR recordings when participants attend to the stimuli and in order to reduce noise present in the recording, FFR researchers encourage their subjects to recline comfortably and/or to sleep during testing (Aiken & Picton, 2008; Skoe & Kraus, 2010). In the present study, participants will be encouraged to sleep or rest quietly during testing.

Subject age. Auditory evoked potentials have shown that neural synchrony in the human auditory system declines with an increase in age (Clinard, Trembaly, & Krishnan, 2010). Clinard et al., (2010) looked at the perceptual and physiological representation of frequency in 32 adults ranging in age from 22-77 years old. Clinard and colleagues used a behavioral discrimination procedure and FFR recordings to examine the representation of frequencies at and around 500 and 1000 Hz as a function of age (2010). Results of the behavioral discrimination procedure showed that the ability to discriminate between different frequencies became poorer as age increased (Clinard et al., 2010). The FFR measures in this study revealed that the neural representation of frequency also declined as age increased, however, these effects were found to be dependent upon the frequency of the stimuli. Differences in phase coherence and amplitude values of the FFR showed that brainstem neural representation of frequencies at and below 1000 Hz was less robust compared to that at frequencies at or below 500 Hz in older adults

(Clinard et al., 2010). These findings are consistent with an age-related decline in neural phase-locking in the brainstem (Clinard et al., 2010).

Similarly, Vander Werff and Burns (2011) recorded FFRs using a synthetic /da/ stimulus from both ears in normal hearing young adults and older adults. Analyses of the sustained FFR components revealed a significant overall age-group difference for all measures. Specifically, the lag time between stimulus presentation and FFR response was significantly longer for the older adult group. Additionally, they found that the spectral magnitude of the harmonic components of the response were significantly smaller for the older adult group. These findings further suggest that the ability of neurons in the brainstem of normal hearing older adults to phase-lock to these components of the stimuli is reduced when compared to younger adults (Vander Werff & Burns, 2011). The present study will attempt to recruit young adults between the ages of 18 and 30 years old.

Gender. Sex differences have been well documented throughout the auditory system. Females have been shown to have shorter ABR wave V latencies and larger wave V amplitudes than males (Jerger & Hall, 1980). Additionally, females are known to have better peripheral hearing sensitivity than males, with males showing greater sound localization abilities than females (Ahadi, Pourbakht, Jafari, Shirjian, and Jafarpisheh, 2014; McFadden, 1998). The FFR has recently been used to examine the sex differences in the subcortical encoding of the speech syllable /da/ in two populations of normal hearing young adults (Ahadi et al., 2014; Krizman, Skoe, & Kraus, 2012b).

Ahadi and colleagues recorded FFRs to binaural presentation of /da/ in 48 young adults, 25 females (2014). They found that females had more robust and better

representation of the fundamental frequency (103-121 Hz), first formant (454-719 Hz), and high frequency components (721-1155 Hz) than males in the spectral domain (Ahadi et al., 2014). Similarly, Krizman, and colleagues analyzed the timing and frequency components of FFRs recorded monaurally to /da/ in 76 young adults, 38 females. They reported sex differences in the spectral encoding of the rapid, higher frequency components of the response (up to 1150Hz), with females having larger spectral magnitudes than males. However, they did not show any significant sex differences with the encoding of the slower, low frequency components of the response such as the fundamental frequency or first formant (Krizman et al., 2012b).

Hoormann, Falkenstein, Hohnsbein, and Blanke (1992) found no significant gender differences for FFR amplitude in their study using tone bursts, however their study included only 20 participants. In an effort to evaluate any potential gender variability seen in the FFR amplitude response, the present study will attempt to recruit an equal number of male and female participants.

Auditory experience. Many researchers have looked at the effects of auditory training, language and music experience on the FFR.

Auditory training. The FFR has been studied pre- and post- auditory training (AT) in various groups of individuals. The findings from these studies have suggested that the subcortical processing of sound is not concrete, but malleable by both short term auditory training and life-long auditory experience with language and music (Russo et al., 2008; Skoe & Kraus, 2010; Song, Skoe, Banai, & Kraus, 2012). Specifically, Song et al. (2012) found that after undergoing auditory training, young adults with normal hearing have a stronger representation of the fundamental frequency (F_0) amplitude in response to

the consonant-vowel (CV) syllable /da/ presented in noise when compared to a control group. Furthermore, Russo et al. (2005) concluded that AT programs have been shown to alter the neural encoding of sound structure at the cortical level, specifically noting the inferior colliculus as the location of neural plasticity (Anderson & Kraus, 2013).

Language experience. The effects of language experience on the FFR have been studied through bilingual language users. Krizman, Marian, Shook, Skoe, and Kraus (2012a) explored whether Spanish/English bilingualism promoted experience-dependent plasticity in the subcortical auditory processing of speech in 48 normal hearing adolescents (25 monolinguals, 23 bilinguals). They found that adolescent bilinguals showed enhanced encoding of the F_0 of the speech syllable /da/ compared to their age-matched monolingual peers (Krizman et al., 2012a). Similarly, Krishnan, Xu, Gandour, and Cariani (2005) studied the representation of pitch through FFRs elicited by 4 Mandarin tones in native speakers of Mandarin Chinese and English. The native Chinese group exhibited stronger pitch representation and smoother pitch tracking in response to the Mandarin tones than the English-speaking group (Krishnan et al., 2005; Krishnan, Gandour, & Bidelman, 2012; Krishnan & Gandour, 2014). Collectively the results of these studies show that bilingualism and tonal language experience promote experience dependent plasticity in the subcortical processing of speech stimuli as evidenced by the FFR (Anderson & Kraus, 2013). This study will recruit native English speakers, without tonal language experience.

Musical experience. Life-long musical experience has also been shown to have effects on the neural representation of speech. Specifically, Parbery-Clark, Skoe, and Kraus (2009) studied the effects of musical experience on the neural representation of

speech in quiet and in noise. When compared to non-musicians, the musicians were found to have a more robust subcortical representation of the speech syllable /da/ in the presence of noise. Furthermore, they found that the neural measures in the musician group were associated with better performance on the behavioral Hearing in Noise Test (HINT) (Parbery-Clark et al., 2009).

Bidelman and Krishnan (2010) examined the effect of one's musical training experience on FFR responses to /i/ presented in quiet and in four different levels of reverberation. There were 10 normal hearing young adults in both the experimental and control groups. The experimental group (musicians) had at least 8 years of consecutive training on a musical instrument and the control group (non-musicians) were required to have less than 3 years of musical training. They found that the musicians had more robust FFRs when compared to the control group in both the quiet test condition and in almost all of the reverberation conditions (Bidelman & Krishnan, 2010).

Parbery-Clark, Anderson, Hittner, and Kraus (2011) studied the benefits of musician-ship on the neural effects of aging. These researchers compared the effects of aging in musicians and non-musicians. They reported that the musicians did not have the age-related neural timing delays in the consonant-vowel transitions that are typically seen in older adults (Anderson & Kraus, 2013; Parbery-Clark et al., 2011). Parbery-Clark and colleagues (2011, 2013) concluded that musical experience limits the negative effects of competition in back ground noise and protects against age-related degradation of neural timing.

In a cross-domain study on the effects of both music and language experience on the representation of pitch in the human auditory brainstem, Bidelman, Gandour, and

Krishnan (2011) recorded FFRs using two stimuli, one from each domain. FFRs were elicited to a musical pitch interval (major third) and a lexical Mandarin tone in native Mandarin speakers, English speaking amateur musicians and English speaking non-musicians. They found that the native Mandarin speakers and musicians had greater pitch-tracking accuracy than the non-musicians. These results provide additional evidence that experience-dependent plasticity can be induced by long-term musical and language experience (Bidelman et al., 2011). The present study will have participants complete a portion of a musical experience survey to quantify each individuals' musical experience.

Stimulus Parameters

The FFR can be recorded to a variety of stimuli types including: simple, complex, time-varying, steady-state, speech, non-speech, consonant-vowel syllables, and tonal sweeps. When the FFR was initially discovered, it was recorded to low frequency tone bursts (Moushegian et al., 1973; Smith et al., 1975; Stillman, Crow, & Moushegian, 1978; Wever & Bray, 1930). More recently, the FFR has been recorded to complex pulse trains (Bidelman, 2015); steady-state vowels such as /a/, /i/, and /u/ (Aiken & Picton, 2008; Ananthkrishnan et al., 2016); time-varying consonant-vowel syllables such as /da/ (Krizman et al., 2012a; Parbery-Clark et al., 2009, 2011; Russo et al., 2005); music (Musacchia, Sams, Skoe, & Kraus, 2007); and Mandarin tones (Krishnan et al., 2005). The stimulus that is chosen for a FFR experiment depends on the population being studied and the research question being asked (Skoe & Kraus, 210). The present study will use simple pure tone stimuli at 250, 500, and 1000 Hz. Simple stimuli at these

frequencies were chosen to eventually compare the responses to the pure tone stimuli that are commonly used clinically.

Stimulus intensity. Skoe and Kraus (2010) recommend a supra-threshold presentation level of approximately 60-85 dB SPL when evoking the FFR. The effects of intensity levels on the FFR have been studied throughout the literature from the early discovery of the FFR in 1973 to recent years (Moushegian et al., 1973). Moushegian et al. (1973) recorded the FFR to a 500 Hz signal in 5 normal hearing individuals at intensity levels ranging from 5 to 56 dB sensation level or SL (i.e., above the subject's behavioral threshold). While their recordings showed a response within 10-20 dB of the subject's behavioral threshold, they noted that a presentation level of at least 46 dB SL is needed to evoke a robust FFR (Moushegian et al., 1973). Similarly, Davis and Hirsh (1976) recorded the FFR to a 500 Hz tone burst in 21 normal hearing individuals at intensity levels ranging from 10 to 70 dB SL in 10 dB increments. They found the threshold of the FFR to be close to 40 dB SL, however, they added that the FFR is rarely clear below a presentation level of 50 dB SL. Their research concluded that the higher presentation levels of 60-70 dB SL produced smoother FFR patterns that were clearly visible (Davis & Hirsh, 1976).

More recent literature has looked at the effects of increasing intensity on the FFR. Krishnan (2002) recorded the FFR to three steady-state vowels at intensity levels between 55 and 85 dB nHL in 10 dB increments. Recordings from 8 normal hearing young adults showed that while the harmonics were represented in the spectra at all intensity levels, the amplitudes of the individual harmonics increased as the intensity of the stimulus increased. In a similar study of FFR recordings to /ba/ at various intensity levels, Akhoun

and colleagues (2008), found that the FFR was best defined at a presentation level of 60 dB SL and was flat at 0 dB SL (the subject's hearing threshold).

As would be expected, FFR recordings from various researchers have shown that response latencies increase as stimulus intensity decreases (Akhoun et al., 2008; Chandrasekaran & Kraus, 2012; Moushegian et al., 1973). As the research outlined above indicates, the clarity and presence of the FFR, as well as the response amplitudes and latency are affected by the intensity of the evoking auditory stimulus. FFR researchers should take care to ensure that the presentation level of their stimuli is consistent across participants as the responses are intensity dependent (Skoe & Kraus, 2010). This study will use a stimulus intensity of 80 dB HL, as it is supra-threshold for listeners with normal hearing thresholds less than 20 dB HL.

Stimulus rate. The stimulus rate is defined as the number of times the stimuli is presented during a set time period, typically one second. The stimulus rate depends on the length of the stimuli and the inter-stimulus interval (ISI), the length of time between the offset of one stimulus and the onset of the next stimulus (Skoe & Kraus, 2010). Skoe and Kraus' (2010) review of the literature has shown that the ISI varies from ~30% of the stimulus length to more than double the length of the stimulus. There are several considerations that need to be made when choosing an ISI. For example, the ISI must be sufficiently long enough so that the response to one stimulus can be completed and return to baseline, before the response to the next stimulus begins. Also, an adequate sample of the baseline activity is needed to evaluate the signal to noise ratio (SNR) of the response (Skoe & Kraus, 2010). As the rate of stimulus presentation increases, with the length of the stimulus (consonant-vowel syllable /da/) fixed at 40 ms, the ISI decreases. A high

presentation rate would have a shorter ISI, giving the response less time to return to the baseline. In contrast, a low presentation rate has a longer ISI, allowing the response ample time to return to baseline before the next stimulus begins.

Krizman, Skoe, and Kraus (2010) recorded the FFR to the consonant-vowel speech syllable /da/ at three different presentation rates; 15.4 Hz (fast), 10.9 Hz (standard), and 6.9 Hz (slow) in 18 normal hearing adults. They found that the FFR was rate dependent for portions of the stimulus greater than 411 Hz, but not for those centered around the F_0 (approximately 103-410 Hz). They proposed that this the high frequency FFR rate dependence suggests an involvement of separate underlying neural mechanisms for high and low frequency FFR components. As the presentation rate increased, the response magnitude decreased for those frequencies greater than 411 Hz (Krizman et al., 2010). The slowest presentation rate (6.9 Hz) showed the greatest magnitude responses for the higher frequency components (Krizman et al., 2010). Based on the evidence above, a slow stimulus rate of 3.13/s will be used in this study.

Stimulus polarity. The polarity of any stimulus (click, tone, speech, etc.) is defined by the initial positive or negative deflection from baseline. Condensation polarities are created through the compression of air molecules; the inward movement of the earphone diaphragm creates an inward movement of the tympanic membrane producing a positive deflection of the basilar membrane from the baseline. In contrast, rarefaction polarities are the decompression of air molecules; the outward movement of the earphone diaphragm creates an outward movement of the tympanic membrane producing a negative deflection of the basilar membrane from the baseline (Kumar, Bhat, D'Costa, Srivastava, & Kalaiah, 2013; Skoe & Kraus, 2010). The FFR can be recorded

using single-stimuli polarity (condensation or rarefaction) or alternating polarities (the stimulus polarity alternates between both rarefaction and condensations polarities) (Kumar et al., 2013). The averaged waveforms from each polarity can then be added or subtracted to accentuate different components of the response, namely the neural envelope or neural temporal fine structure (Aiken & Picton, 2008; Ananthakrishnan et al., 2016; Skoe & Kraus, 2010). This process as it relates to analyzing the FFR will be more thoroughly explained in the “What information does the FFR provide us” section of this literature review. This study will record FFRs to an alternating polarity stimulus.

Recording Parameters

Electrode montage. The FFR is typically recorded using a one-channel vertical montage, however, two-channel vertical and horizontal recording montages are also used. Bidelman (2015)’s multichannel analysis of FFR scalp topography confirmed that the FFR is maximally recorded over the frontal midline of the scalp, suggesting that the FFR is optimally recorded using Fz or Fpz as the active, non-inverting electrode. These findings are consistent with what is currently used throughout the FFR literature. Skoe and Kraus (2010) report using a single channel recording with electrodes placed at Cz (active), ipsilateral ear lobe (reference) and forehead as ground for rostral brainstem recordings throughout their research studies. Other researchers have used a two-channel vertical and horizontal FFR recording montage. In their two-part experiment, Aiken and Picton (2008) used a vertical montage from Cz (vertex) to C7 (mid posterior neck) and the horizontal montage recorded between the left and right mastoids. Furthermore, two different vertical montages can be used in a two-channel recording and then averaged together to create a response with a higher signal to noise ratio (SNR). Ananthakrishnan

et al. (2016) report using a two-channel recording with the active electrode at the midline of the forehead at the hair line (Fz) for both channels, referenced to linked mastoids (A1/A2) for channel 1 and then to the 7th cervical vertebra (C7) in channel 2, with a mid-forehead (Fpz) electrode as the common ground. Bidelman (2015) concluded that the amplitude of the FFR depends on the electrode montage that is used and the orientation of the electrodes to the neural generators of the recordable response. Based on the research outlined above this study will use a two channel vertical electrode montage.

Sampling rate. The sampling rate, or digitization rate, of an FFR recording refers to the number of times the neural signal is digitally sampled by the recording system in one second. Skoe and Kraus (2010) recommend a sampling rate of 6000-20,000 Hz for optimal temporal precision. The FFR requires a high sampling rate to accurately digitize the fast, phase-locked components that are recorded from the brainstem (Bidelman, 2015). The Nyquist theorem suggests a sampling rate that is twice the highest frequency in the stimulus, however most FFR researchers sample well above the Nyquist frequency (Krishnan et al., 2005; Russo et al., 2005; Skoe & Kraus, 2010). A high sampling rate not only reduces sampling errors, but also increases the temporal precision of the FFR recording and allows for the response peaks to be more clearly recognized (Skoe & Kraus, 2010). A sampling rate of 20,000 Hz will be used for this study.

Analysis window. The analysis or averaging window is the time period in seconds during which the neural signals are averaged. The averaging window should be sufficiently long to include a 10-50 ms pre-stimulus period, the length of the stimulus and response period, and a 10-50 ms post-stimulus period. The pre-stimulus period is

necessary in calculating the peak amplitude of the response by allowing for the calculation of the ambient EEG amplitudes (noise floor), against which the response peak amplitudes are compared. The post-stimulus time period allows the neural activity to return to baseline prior to the start of the next analysis (Skoe & Kraus, 2010). The analysis window for this study will be 300 ms.

Sweeps. The number of sweeps, individual stimulus presentations, averaged over the allotted time frame determines the robustness and reliability the FFR response. Skoe and Kraus (2010) recommend averaging approximately 2000-3000 sweeps per polarity for a total of 4000 to 6000 total sweeps. The signal to noise ratio of the response is directly proportional to the square root of the number of sweeps in the response, thus as you increase the number of sweeps the SNR increases also (Skoe & Kraus, 2010). This study will be recorded using 2000 sweeps per stimuli.

What Does the FFR Tell Us?

The FFR reveals how well the human auditory brainstem encodes the acoustical characteristics contained within the eliciting stimulus. Specifically, the FFR elicited by a complex stimulus in different stimulus polarities, can be manipulated in order to measure brainstem phase-locking ability to the neural envelope (FFR_{ENV}), conveying information about low frequency components of the stimulus such as F_0 , and to the neural temporal fine structure or TFS (FFR_{TFS}), which conveys information about high frequency components of stimulus such as formant frequencies. The FFR reflects the neural encoding of both FFR_{ENV} and FFR_{TFS} for lower frequency stimuli. However, as the stimulus frequency approaches the limits of neural phase-locking in the inferior

colliculus, 2000 Hz, the FFR reflects a stronger encoding of the FFR_{ENV} rather than the FFR_{TFS} (Greenberg et al., 1987; Moushegian et al., 1973, Skoe & Kraus, 2010).

As previously discussed in the Stimulus Polarity section above, the FFR can be recorded using a single-polarity stimulus (condensation or rarefaction) or an alternating polarity stimulus (the stimulus polarity alternates between both rarefaction and condensation polarities) (Kumar et al., 2013). The averaged waveforms from each polarity can then be added or subtracted to accentuate brainstem representation of different components of the response, namely the envelope (FFR_{ENV}) or temporal fine structure (TFS) (FFR_{TFS}) (Aiken & Picton, 2008; Ananthakrishnan et al., 2016; Skoe & Kraus, 2010). If collected in alternating polarity, the response must first be separated into its condensation and rarefaction components prior to extracting the FFR to the envelope and TFS separately. The Intelligent Hearing Systems' (IHS) Smart Evoked Potential software is able to separate the alternating polarity waveform through its "split buffer averaging" feature. When the desired waveform is selected, this feature will split the waveform into the separate condensation and rarefaction polarities. Figure 9 illustrates this process with an FFR time waveform recorded to the complex vowel sound /u/.

When the condensation and rarefaction waveforms are in their individual forms they are then either added or subtracted to extract the envelope or TFS of the FFR response. The addition of the condensation and rarefaction polarity waveforms creates a summed FFR waveform that is primarily phase-locked to the neural envelope, reflecting low-frequency stimulus content (Skoe & Kraus, 2010). In contrast, the subtraction of the responses to the condensation and rarefaction polarities creates a FFR waveform that is primarily phase-locked to the TFS or higher frequency components of the stimulus (Skoe

& Kraus, 2010). See Figure 10 for examples of an envelope and TFS temporal waveform.

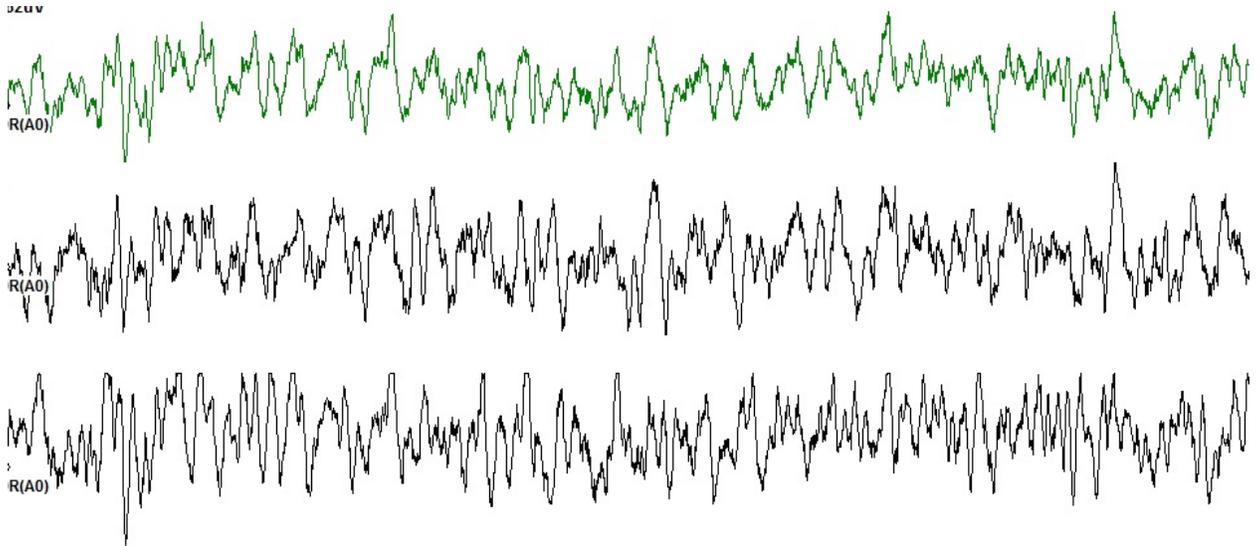


Figure 9. The green waveform is the original FFR recording to a steady-state English back vowel /u/ with an $F_0=120$ Hz and an $F_1=360$ Hz. The waveform was then split into the two polarities (condensation and rarefaction) using split buffer averaging in the IHS SmartEP data acquisition software.

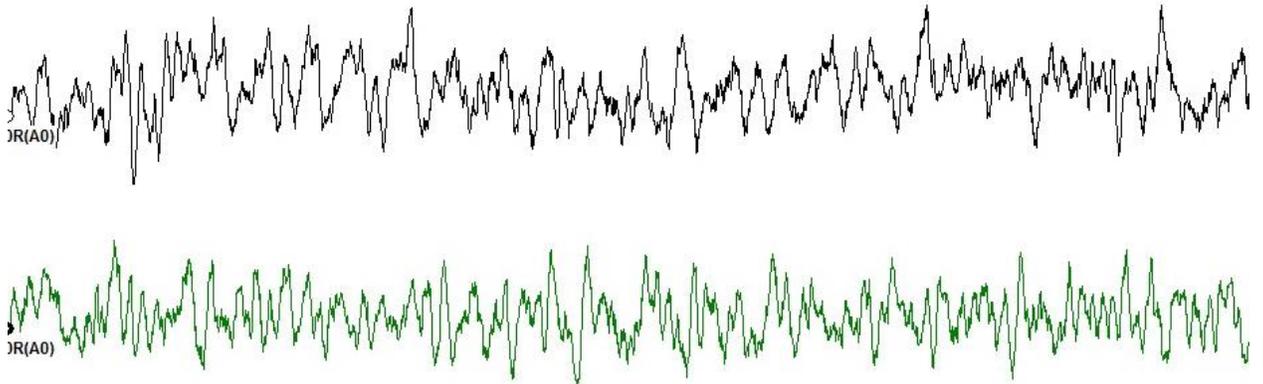


Figure 10. The black waveform is the Envelope FFR, resulting from the addition of the responses to the condensation and rarefaction polarity stimuli. The green waveform is the TFS FFR, resulting from the subtracted response of the condensation and rarefaction polarities.

In order to determine the frequency content in these time waveforms, they must first be converted into the frequency domain using a Fast Fourier Transform (FFT) analysis. As previously discussed, the FFT allows us to view the energy (amplitude) of individual frequency components making up the temporal waveforms (Emanuel & Letowski, 2009).

When an FFT is used to break down a complex waveform such as FFR_{ENV} recorded in response to a complex stimulus, the spectral components reveal the greatest amount of energy at the fundamental frequency or F_0 of the eliciting stimulus (see Figure 11). For example, consider an eliciting stimulus like a complex speech sound such as the vowel /u/ with an F_0 of 120 Hz and an F_1 of 400 Hz. The FFT of the FFR_{ENV} recorded for this stimulus will show a robust spectral peak at approximately 120 Hz, as is seen in Figure 11. On the other hand, an FFT of the FFR_{TFS} reveals frequency energy at the formant-related harmonics of the F_0 that are present in the waveform (See figure 12).

When an FFR is recorded to a pure tone stimulus, the FFT of the FFR_{ENV} provides an amplitude response at a frequency that is exactly double what was used to evoke the response, whereas, the FFT of the FFR_{TFS} give a response at the same frequency that was used to evoke the response. Therefore, the present study will only look at the FFR_{TFS} of the recorded response.

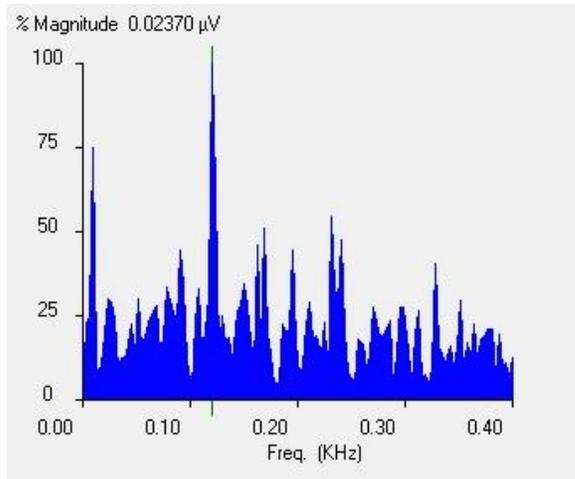


Figure 11. FFT of the FFR_{ENV} (black temporal waveform in Figure 10). This FFT reveals a primary spectral component with a large magnitude at 120 Hz, the fundamental frequency of the vowel /u/.

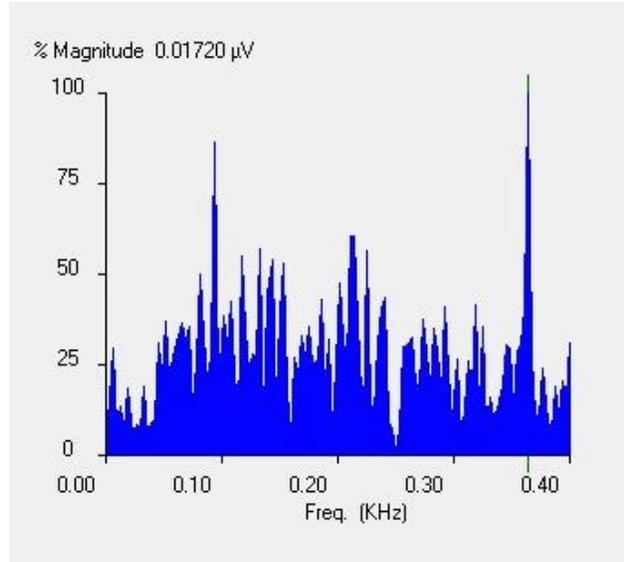


Figure 12. FFT of the FFR_{TFS} (green temporal waveform in Figure 10). This FFT reveals a primary spectral component with a large magnitude at approximately 400 Hz, the first formant frequency of the vowel /u/.

Research Question

Despite the FFR being a relatively young response (~40 years), as compared to some of the other evoked potentials, researchers have learned a lot about this response in a relatively short amount of time. However, one aspect of the FFR that has proven to be hard to explain is the significant amount of variability seen in this response, even within a group of normal hearing listeners (Ruggles, Bharadwaj, & Shinn-Cunningham, 2011). Such response variability has been attributed to a number of factors including but not limited to experience or training induced plasticity, gender and age of the subject, “hidden hearing loss” (normal audiometric thresholds with difficulties in supra-threshold speech perception), etc.

Through years of research, the FFR has been shown to be a sensitive and reliable measure of experience-dependent plasticity with musical experience, auditory training, and language experience (Krishnan et al., 2012; Krizman et al., 2012a; Russo et al., 2008; Skoe & Kraus, 2010). Perhaps one of the most widely studied aspects of brainstem neural plasticity is that consequent to music training. As discussed earlier, multiple studies have established stronger FFRs in musicians when compared to non-musicians (Anderson & Kraus, 2013; Bidelman & Krishnan, 2010; Parbery-Clark et al., 2011). However, once again, such training/experience dependent neural plasticity has been observed in response to certain stimuli that mimic music contours, and it remains unclear if these training effects can be observed for FFRs elicited by pure tone stimuli.

Additionally, the influence of gender on FFR amplitude has been a source of variability in the FFR literature (Ahadi et al, 2014; Krizman et al., 2012b). In general, females have been shown to have greater neural encoding of the F_0 and first formant (F_1)

(Ahadi et al, 2014,); and high frequency components (Krizman et al., 2012b) in response to the speech syllable /da/. While some studies using tonal stimuli have shown no effects of gender on the FFR amplitude (Batra et al., 1986; Hoormann et al., 1992). It is unknown whether gender differences will be evident in FFRs elicited by pure tone stimuli.

Further, there are also other sources of variability: stimulus-based, and analysis technique-based. As has been discussed previously, the FFR can be recorded to varying stimuli, in multiple populations, and analyzed using different techniques depending on the research question. More specifically, the FFR can be recorded to a variety of stimuli types including: simple, complex, time-varying, steady-state, speech, non-speech, consonant-vowel syllables, and tonal sweeps. Given that the FFR is a direct reflection of the stimulus, it stands to reason that as the stimuli become more complex, the resultant FFR, in turn, is more complex as well. Further, researchers use a number of different algorithms to measure the strength of brainstem neural encoding of pitch reflected by the FFR. These analysis techniques range from Fast Fourier Transform (FFT) analyses and autocorrelation, which look at the strength of phase-locked activity to the F_0 and amplitude envelope, to sliding window analyses which look at how the FFR changes over time (Skoe & Kraus, 2010).

Lastly, the variety of stimuli and analysis techniques used across research studies makes it difficult to understand the variable nature of the complex response that is the FFR. A first step that may help researchers to better understand the subject-related variability reflected in the brainstem response could be to minimize stimulus-related

complexity by recording the FFR in response to simple stimuli such as pure tones, which are also used clinically, in a group of normal hearing adults.

Hence, the specific objectives of this study are twofold: 1) to examine the effect of frequency on the FFR to low frequency pure tone stimuli (250, 500 and 1000 Hz) in a group of normal-hearing individuals, and 2) to examine subject-related variability in brainstem neural encoding strength, specifically, if and how gender and varying degrees of musical experience are reflected in the neural encoding of a 500 Hz pure tone stimuli.

CHAPTER 2: METHODOLOGY

Participants

A total of 30 young adults (male =12, age range: 19-29 years, M=23.71) participated this study. All of the participants were native English speakers with normal hearing sensitivity, and audiometric thresholds at or below 15 dB HL from 250-8000 Hz. All participants had normal middle ear function in the test ear, evidenced by a type A tympanogram. Informed consent was obtained from all participants prior to participation in the study in accordance with the Institutional Review Board of Towson University. All participants were compensated with a \$20 gift card to Starbucks or Target.

Materials and Measures

Stimuli. FFRs were recorded using 250, 500 and 1000 Hz pure tone stimuli, as they are the stimulus frequencies tested regularly in clinical practice and within the limits of phase locking in the auditory brainstem. The stimuli were presented at an intensity of 80 dB HL. The duration of each stimulus was 250 ms.

Procedure. Audiometric test results were obtained from 250-8000 Hz for each participant prior to FFR recordings using a GSI-61 audiometer calibrated on August 11, 2016. Additionally, participants filled out page 5 of the Goldsmith's Musical Sophistication Index (Müllensiefen, Gingras, Musil, & Stewart, 2014), and a case history questionnaire. The FFR was recorded with the participants reclined and relaxed in an acoustically and electrically shielded darkened sound booth. Participants were encouraged to sleep during the FFR recording. Conductive paste filled disc electrodes were fixed to disinfected skin abraded with a skin prepping gel with electrode impedances less than 5 k Ω . The FFR was recorded using a two-channel recording with

the active electrode placed at the midline of the forehead at the hair line (Fz) for both channels, referenced to linked mastoids (A1/A2) for channel 1 and then to the 7th cervical vertebra (C7) in channel 2, with a mid-forehead (Fpz) electrode as the common ground. IHS Smart EP version: 3.92 (Intelligent Hearing Systems, FL, USA), was used to record the FFR.

The stimuli were delivered monaurally to the right ear through electronically shielded Etymotic Research ER-3 earphones in alternating polarities at a rate of 3.13/sec over a 300 ms analysis window. Each recording collected 2000 sweeps. Each recording was repeated twice to ensure waveform replicability. A total of 4000 sweeps were analyzed per participant per stimulus frequency tested. The presentation order of stimuli was randomized across participants. The full experimental protocol took approximately 90 minutes per participant.

Data analysis. MATLAB, a data processing software, was used to convert each participants' IHS time waveform data files into .wav files. The waveforms were then split into condensation and rarefaction polarities. The condensation and rarefaction polarity responses were added and subtracted to extract the FFR_{ENV} and FFR_{TFS} , respectively. Temporal FFR waveforms were spectrally analyzed through a Fast Fourier Transform (FFT) and the magnitude of the spectral peaks of the FFR_{TFS} was recorded per condition, per subject.

Participants were split into three groups based on their self-reported number of years of formal musical training (0, 1-5, and 6-10+) as reported on the Goldsmith's Musical Sophistication Index (Müllensiefen et al., 2014).

Statistical analysis. A repeated measure analysis of variance (ANOVA) was used to measure the effect of pure tone stimulus frequency (250, 500 and 1000 Hz) on the magnitude of the FFR_{TFS}. Additionally, a two-way ANOVA was conducted to examine the effect of gender and musical training on FFR_{TFS} FFT magnitude at 500 Hz.

This study looked at multiple independent variables including: gender, musical ability (number of years of formal musical training), and frequency of the pure tone eliciting the FFR (250, 500 and 1000 Hz). The dependent variable was the magnitude of the FFR F_0 . SPSS was used for all statistical analyses.

CHAPTER 3: RESULTS

Figure 13 shows 150ms of the grand averaged FFR_{TFS} waveforms, (overlaid for the three stimuli tested, 250, 500 and 1000 Hz) across the 30 participants. The FFR waveforms reflect the subtracted temporal waveforms, revealing the temporal fine structure encoding, FFR_{TFS} . The FFR temporal waveform amplitudes appear to increase with decreasing stimulus frequency, with smaller amplitudes seen for 1000 Hz than for 500 and 250 Hz stimuli. The periodicity of both the 250 and 500 Hz waveforms appear more uniform than the 1000 Hz waveform. The grand averaged FFR_{TFS} FFT magnitude, plotted as a function of pure tone stimulus frequency at 250, 500 and 1000 Hz for the 30 participants can be seen in Figures 14 and 15. Figures 14 and 15 reinforce the observations in the temporal waveform data seen in Figure 13. Specifically, the average FFR_{TFS} FFT magnitude increases as the frequency of the stimulus decreases.

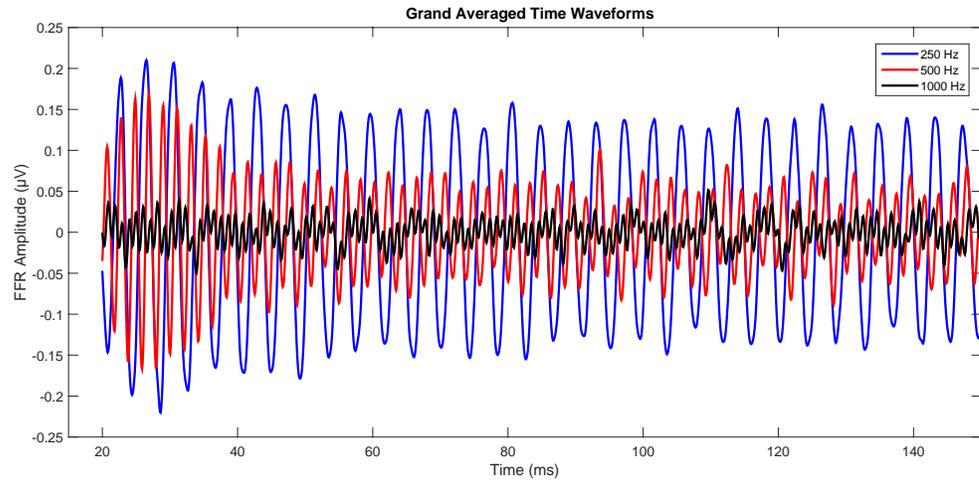


Figure 13. Grand Averaged FFR_{TFS} waveforms for the three stimulus frequencies tested plotted as a function of time (ms). These waveforms reflect the subtracted condensation and rarefaction waveforms, revealing the temporal fine structure encoding, FFR_{TFS} .

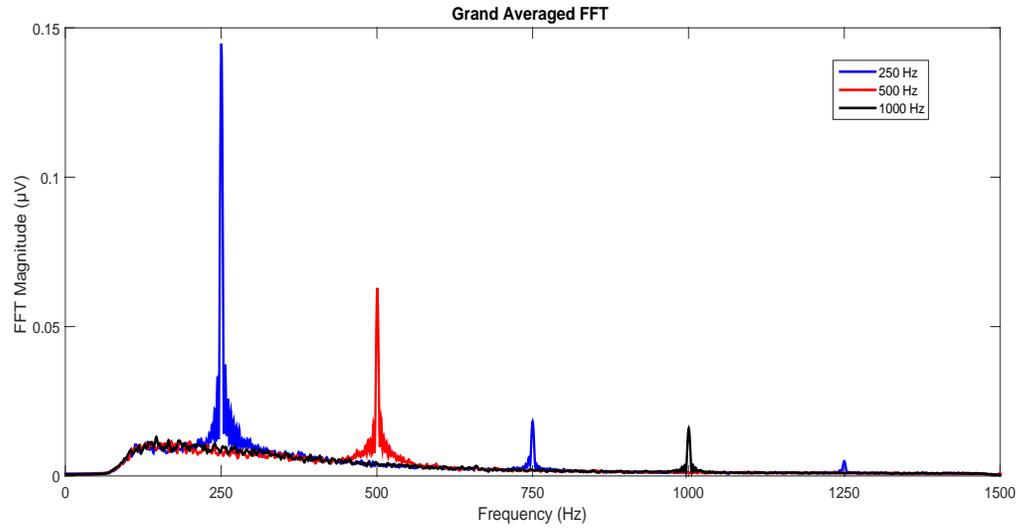


Figure 14. This figure shows the grand averaged FFRTFS FFT magnitude measured in μV , as a function of frequency in Hz. The peak FFTs for each frequency line up with the frequency of the pure tone stimuli used in the FFR recording.

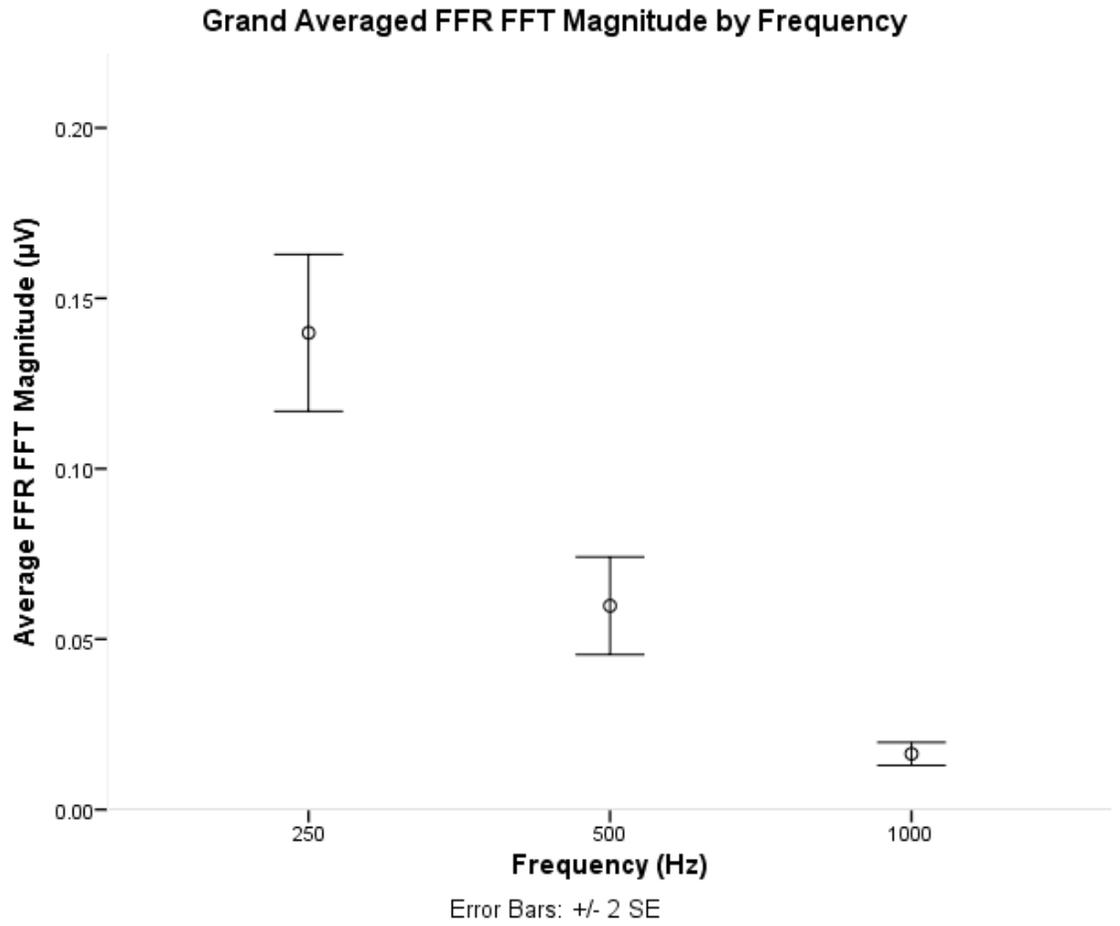


Figure 15. This graph shows the grand averaged FFR_{TFS} FFT magnitude measured in µV, plotted as a function of pure tone stimulus frequency in Hz. The error bars represent standard error.

A repeated measures ANOVA was performed on the FFR_{TFS} data obtained at the three different stimulus frequencies (250, 500 and 1000 Hz). Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 10.559, p = 0.005$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .761$). The results show that there was a significant main effect of stimulus frequency on FFR_{TFS} magnitude $F(1.522, 44.135) = 76.2, p < .001$.

Bonferroni corrected ($\alpha = 0.0167$) post hoc tests indicated that there was a significant difference between the average FFR_{TFS} magnitude of the 250 Hz stimulus and both the 500 and 1000 Hz stimuli ($p < .001$). Additionally, there was a significant difference between the average FFT amplitude of the 500 Hz stimulus and the 1000 Hz stimulus, $p < .001$.

As previously mentioned, participants were split into three groups based on their self-reported number of years of formal musical training (0, 1-5, and 6-10+). A two-way ANOVA was conducted to examine the effect of gender and musical training on FFR_{TFS} magnitude at 500 Hz. There was a significant main effect of gender on FFR_{TFS} magnitude $F(1, 24) = 7.103, p < .05$, indicating that females had larger FFR_{TFS} magnitudes than males at 500 Hz. There was a non-significant main effect of musical training, $F(2, 24) = 3.245, p = .057$. Additionally, there was a non-significant interaction between gender and years of musical training $F(2, 24) = 1.562, p = .230$.

Figure 16 shows a graphical representation of the average FFR_{TFS} magnitude at 500 Hz, plotted as a function of gender and years of musical training for the 30 participants.

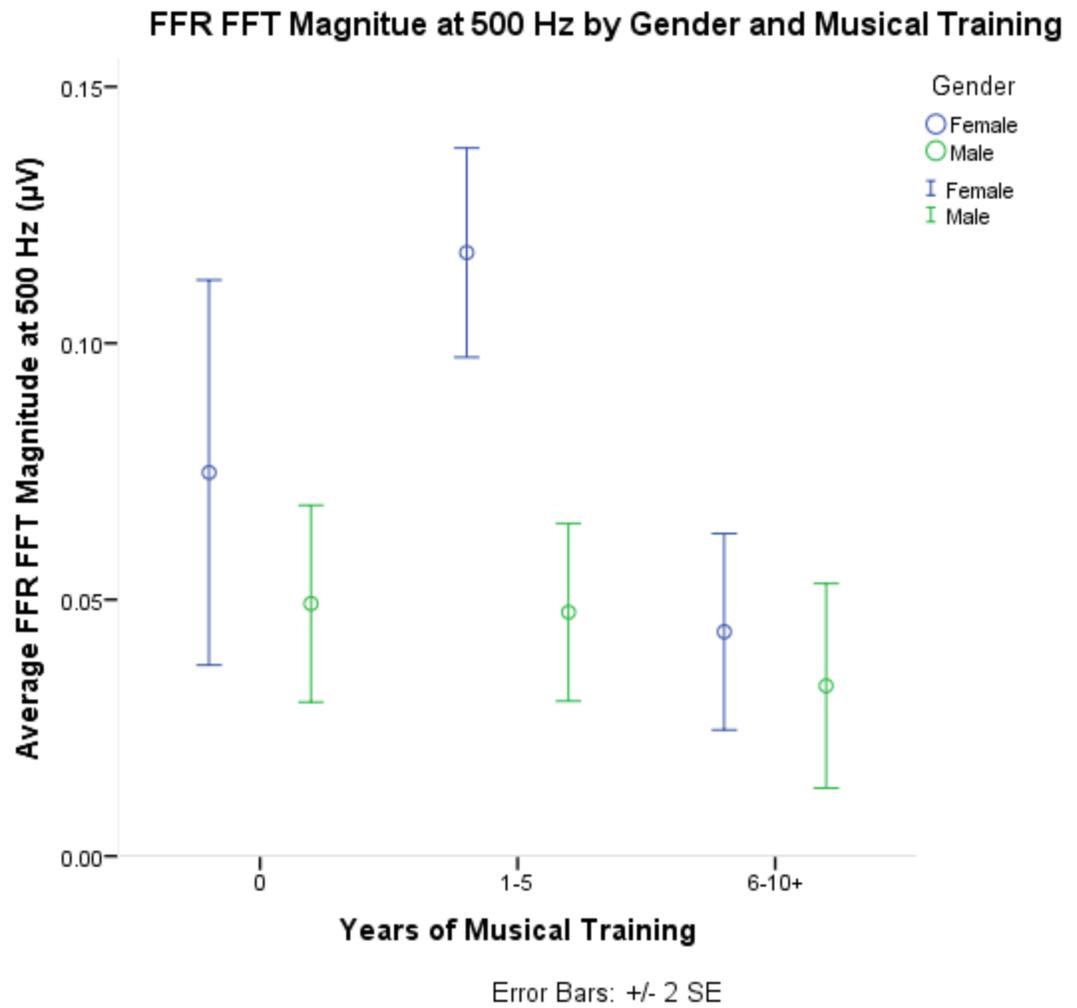


Figure 16. This graph illustrates the average FFR_{TFS} FFT magnitude at 500 Hz measured in μV , plotted as a function of gender and number of years of musical training. The error bars represent standard error.

CHAPTER 4: DISCUSSION

The objectives of the present study were twofold: the first was to examine the effect of stimulus frequency on the FFR to low frequency pure tone stimuli (250, 500 and 1000 Hz) in a group of normal-hearing individuals, and second to examine the subject-related variability in brainstem neural encoding strength of a 500 Hz pure tone stimuli.

Effects of Frequency on FFR Amplitude

The present study revealed a significant main effect of frequency on brainstem neural representation of pitch. Specifically, the average FFR_{TFS} magnitude at 250 Hz was significantly larger than both the 500 and 1000 Hz, and 500 Hz magnitude was significantly larger than the 1000 Hz magnitude. These findings suggest that neural phase locking strength is greater for low frequency pure tone stimuli, when compared to frequencies, such as 1000 Hz, that are closer to the upper limits of phase locking in the human auditory brainstem, and are consistent with previous literature (Batra, Kuwada, & Maher, 1986; Hoormann et al., 1992; Moushegian et al., 1973; Rose et al., 1967; Skoe & Kraus, 2010).

Effects of Gender on FFR Amplitude

This study revealed a significant main effect of gender on brainstem neural representation of 500 Hz, as indexed by the FFT magnitude of the FFR_{TFS} at 500 Hz. As a group, females had larger FFR_{TFS} FFT magnitudes than males at 500 Hz. The variability among FFR recordings in males and females has been studied throughout the literature.

As was previously mentioned, the FFR has recently been used to examine the sex differences in the subcortical encoding of the speech syllable /da/ in two populations of

normal hearing young adults (Ahadi et al., 2014; Krizman, Skoe, & Kraus, 2012b).

Ahadi and colleagues found that females had more robust and better representation of the fundamental frequency (103-121 Hz), first formant (454-719 Hz), and high frequency components (721-1155 Hz) than males in the spectral domain (2014). Similarly, Krizman, and colleagues reported sex differences in the spectral encoding of the rapid, higher frequency components of the response (up to 1150Hz), with females having larger spectral magnitudes than males (2012b). However, they did not show any significant sex differences with the encoding of the slower, low frequency components of the response such as the fundamental frequency or first formant (Krizman et al., 2012b).

The findings of Ahadi et al., (2014) do not entirely support the initial findings of Krizman et al. (2012b), however, the researchers proposed that the differences could be due to the different stimulation modes that were used (binaural vs. monaural). Additionally, Ahadi et al., (2014) used a significance value of $p < .05$, while Krizman et al., (2012b) used a significance value of $p < .001$. With this difference taken into account, the spectral magnitude of the first formant frequency range in the Krizman et al. (2012b) paper would be significantly larger for females than for males at the $p < .05$ level.

Despite the use of pure tone stimuli as opposed to the speech syllable /da/ the findings of the present study are consistent with the literature presented above (Ahadi et al., 2014; Krizman et al., 2012b). For both the studies presented above, using a significance value of $p < .05$, there was a significant difference between males and females in the spectral representation of first formant range of the speech syllable /da/ which falls in the 220-720 Hz range (Krizman et al., 2012b).

In contrast, Hoormann and colleagues (1992) found no significant gender differences for FFR amplitude in their study using tone bursts. However, when compared to the two studies described above, Hoormann et al., (1992), had a significantly less number of participants, $N=20$, where Ahadi et al., (2014) and Krizman et al., (2012b) had 48 and 76 participants, respectively.

The present study reported a significant sex difference in the magnitude of the FFR_{TFS} FFT response at 500 Hz, with females showing larger responses than males. Given that a 500 Hz pure tone falls within the frequency range of the first formant of the speech syllable /da/, these results appear consistent with the findings presented by Hoormann et al. (2014) and Krizman et al. (2012b). At first glance, these results suggest that females may exhibit stronger brainstem neural synchrony than males. In FFR research, gender differences are seen in the neural encoding of the fast, transient elements of speech, as opposed to the slow, sustained elements (Ahadi et al., 2014; Krizman et al., 2012b). Various theories have been presented to describe driving factor(s) behind the gender differences seen in the literature including: smaller head size, female sex hormones, and shorter cochlea length (Ahadi et al., 2014; Krizman et al., 2012b). However, the current findings should be interpreted with caution, especially given the low sample sizes involved. The present study had 12 male and 18 female participants, the difference in number between the two groups may have affected the statistical analyses that were completed.

Effects of Musical Training on FFR Amplitude

While previous research on the FFR and formal musical training have shown significant differences in the subcortical representation of speech, the present study did not show a significant effect of musical training on the FFT amplitude of the FFR_{TFS} recorded to pure tone stimuli. The differences between the present results and those of previous studies could potentially be explained by differences in the stimuli used to elicit the FFR, the classifications used to quantify an individual's musical training, and low sample size.

The present study looked at simple stimuli, pure tones, that are a poor representation of the behaviorally relevant sounds that one encounters in daily life (Skoe & Kraus, 2010). Complex stimuli are better approximations of the sounds that are regularly encountered outside the audiology booth, such as speech, music, and environmental sounds. Recent studies that have evaluated the effects of musical training on the FFR amplitude utilized complex stimuli such as, native speech sounds (/da/, /i/), musical pitch intervals, and linguistic pitch contours to record the response (Bidelman et al., 2011; Bidelman and Krishnan, 2010; Chandrasekaran & Kraus, 2010; Musacchia et al., 2007; Parbery-Clark et al., 2009; Skoe, Burakiewicz, Figueiredo, & Hardin, 2017). The differences in stimuli used could have influenced the strength of the neural encoding that was observed in the present study, as well as prevented a significant difference in FFT response amplitude among the varying levels of musical experience.

Alternatively, researchers that have specifically recruited for musicians and non-musicians, have used different criteria for classifying a musician, with most defining a non-musician as someone with less than 3 years of musical experience. Parbery-Clark

and colleagues (2009) classified musicians as individuals who started instrumental training before the age of 7 years and who had practiced consistently for at least 10 years prior to participation in the study, non-musicians could not have more than 3 years of musical experience and the experience had to be more than 7 years before participation in the study. As previously mentioned, Parbery-Clark et al. (2009) found that when compared to non-musicians, musicians had a greater subcortical representation of the speech syllable /da/ in the presence of noise.

Bidelman and Krishnan (2010) examined the effect of one's musical training experience on FFR responses to /i/ presented in quiet and in four different levels of reverberation. Their experimental group (musicians) had at least 8 years of consecutive training on a musical instrument and the control group (non-musicians) were required to have less than 3 years of musical training. They concluded that the musicians had more robust FFRs when compared to the control group in both the quiet test condition and in almost all of the reverberation conditions (Bidelman & Krishnan, 2010).

Bidelman et al., (2011) recorded FFRs to a musical pitch interval (major third) and a lexical Mandarin tone in native Mandarin speakers, English speaking amateur musicians and English speaking non-musicians. They found that the native Mandarin speakers and musicians had greater pitch-tracking accuracy than the non-musicians. A musician in this study was defined as an instrumentalist with at least 9 years of continuous training in western classical music, beginning before age 11 years, who currently played their instruments and had formal training in the 5 years prior to their participation in the study. The participants in the non-musician and Mandarin speaking

groups could not have had more than 3 years of formal musical training (Bidelman et al., 2011).

Not only did the present study differ from the literature in the use of simple vs. complex stimuli, participants were not specifically recruited based on previous musical training. As previously mentioned, participants were asked to complete page 5 of the Goldsmith Musician Sophistication Index and were split into three groups based on their self-reported number of years of formal musical training (0, 1-5, and 6-10+). However, if participants were intentionally recruited for grouping of musicians and non-musicians, there is not a commonly accepted method for quantifying ones' past musical experience. The Goldsmith Musician Sophistication Index in its entirety could be a prospective tool for quantifying musicianship among normal hearing listeners, although it is quite lengthy and the scoring is very detailed. There remains a need for a simple, efficacious, and standard quantification measure of musical ability for use in the literature. The differences in stimuli and quantification of musician training between the present study and those in the literature could justify the absence of a significant effect of musical training on FFR_{TFS} FFT amplitude. Additionally, the present study had a small sample size. Once the participants were grouped according to number of years of formal musical experience, the group sizes were as follows: 0 years (N=13, females = 8), 1-5 years (N=7, females = 3), 6-10+ years (N=10, females = 7). As with the gender statistical analyses, the difference in number of participants within the groups may have affected the statistical analyses that were completed.

Clinical Implications

The variety of stimuli and analysis techniques used across research studies makes it difficult to understand the variable nature of the complex response that is the FFR. The inherent subject-related variability of the FFR coupled with external sources of variability (the wide range of stimuli, populations, and analysis techniques used by different researchers when recording the FFR) render deciphering the complex nature of the FFR challenging. The present study attempted to better understand the nature of this complex response by minimizing the sources of “external” variability, by measuring the response to simple stimuli, such as pure tones, in a group of normal-hearing young adults. Since pure tone stimuli are regularly used in the audiological test suite, the results from this study can be used to provide a better understanding of the neural encoding of the stimuli that are commonly used in everyday audiological testing. However, before the FFR can be used as a clinical tool, further research is needed to further define sources of subject-related variability affecting the FFR and its recording.

Study Limitations and Future Directions

The present study had several limitations. One of the main limitations was the lack of musical diversity among our participants. Significant efforts were made to recruit persons from outside of the audiology department, however, we did not specifically recruit for experimental groups of musicians and non-musician controls. Also, we simply quantified one’s musical ability into three groups based on the number of self-reported years of formal musical training, a single question from the Goldsmith Musician Sophistication Index. The use of a more formal quantification measure such as the entirety of the Goldsmith Musician Sophistication Index may have provided more

cohesive groups with less variability among them. Additionally, the current data set reflects a small sample size. The relationship between gender and musical ability on the brainstem neural representation of frequency would be better understood by increasing the number of subjects and specifically recruiting for musicians and non-musicians.

There are many subject-related variables that can affect the strength of the FFR response including: experience or training induced plasticity, gender and age of the subject, “hidden hearing loss” (normal audiometric thresholds with difficulties in supra-threshold speech perception), etc. In an attempt to examine the variability of the FFR, we observed a great variability within the number of years of formal musical training that the participants reported.

Future studies may wish to examine the variability of the FFR recorded to pure tone stimuli among a wider range of variables including: pure tone thresholds, noise exposure, linguistic status, speech recognition abilities and performance on a range of auditory processing tests. Researchers may also wish to quantify years of musical training to serve as a covariate in the analysis of the FFRs (Skoe et al., 2017).

APPENDIX A: Informed Consent Form

Project title: Frequency Following Response (FFR): Examining the Variability of this Complex Response

Principal Investigators:

Lindsay Roberts B.A.
lrober11@students.towson.edu

Laura Grinstead, B.S.
lgrins1@students.towson.edu

The objective of this thesis project is to record Frequency Following Responses (FFR) to pure-tone stimuli on normal hearing individuals. All participants will be native speakers of English between the ages of 18 and 30 years old. The FFR is an electrophysiological neural response (EEG-type) that is recorded from sensors placed on the scalp. The FFR is able to provide information regarding the neural encoding of pitch in the human auditory brainstem when sounds are presented to the ear. We are going to record FFRs to pure tone stimuli at 250, 500, and 1000 Hz and record the magnitude of the recorded response. Additionally, we will look at the musical experience of the participants to see if musical experience is reflected in the strength of the response. Your role in this project will consist of attending two sessions: one 30-minute session for hearing test, and one 90-minute session for FFR recording.

At the test session, you will be asked to answer some questions about your hearing and musical experience history, have your hearing screened, and have an FFR recorded. The session itself consists of the test administrator placing 5 sensors (electrodes) over the surface of the head, behind both ears, and back of neck. A soft foam ear plug will be placed in your right ear canal. You will then be instructed to get comfortable and relax in a reclining chair while listening to sounds presented to your ear. There are no known benefits, risks or discomforts associated with this procedure.

Participation in this study is voluntary. All information will remain strictly confidential. Although the descriptions and findings may be published, at no time will your name be used. All data obtained electronically will be password protected and any personal identification will be removed. All paper information collected during the study will be kept strictly confidential and will be located in a locked cabinet. You are at liberty to withdraw your consent to the experiment and discontinue participation at any time without prejudice. If you have any questions after today, please feel free to call Dr. Ananthakrishnan, the faculty advisor at 410-704-6369 or contact Dr. Elizabeth Katz, Chairperson of the Institutional Review Board for the Protection of Human Participants at Towson University at (410)-704-2236.

I, _____, affirm that I have read and understood the above statement and have had all of my questions answered.

Signature: _____ Date: _____
Witness: _____ Date: _____

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

APPENDIX B: Case History Form

Patient ID: _____

- Have you ever had your hearing tested before? If Yes, what were the results?
- Do you feel like you have hearing loss?
- Does anyone in your family have hearing loss?
- Do you have any tinnitus? (ringing, buzzing or roaring in your ears)
- Do you ever experience dizziness, balance problems, or falls?
- Do you have any ear pain?
- Have you had any ear infections as an adult?
- Have you had any ear drainage in the last 30 days?
- Have you ever had ear surgery?
- Have you ever been exposed to recreational or occupational noise (ex. Military, factory, music, gun fire?).
- Have you had formal musical training? If so, how many years? What instruments?

APPENDIX C: Goldsmith Musician Sophistication Index (Page 5)

Please circle the most appropriate category:

32. I engaged in regular, daily practice of a musical instrument (including voice) for 0 / 1 / 2 / 3 / 4-5 / 6-9 / 10 or more years.
33. At the peak of my interest, I practiced 0 / 0.5 / 1 / 1.5 / 2 / 3-4 / 5 or more hours per day on my primary instrument.
34. I have attended 0 / 1 / 2 / 3 / 4-6 / 7-10 / 11 or more live music events as an audience member in the past twelve months.
35. I have had formal training in music theory for 0 / 0.5 / 1 / 2 / 3 / 4-6 / 7 or more years.
36. I have had 0 / 0.5 / 1 / 2 / 3-5 / 6-9 / 10 or more years of formal training on a musical instrument (including voice) during my lifetime.
37. I can play 0 / 1 / 2 / 3 / 4 / 5 / 6 or more musical instruments.
38. I listen attentively to music for 0-15 min / 15-30 min / 30-60 min / 60-90 min / 2 hrs / 2-3 hrs / 4 hrs or more per day.
39. The instrument I play best (including voice) is ____

APPENDIX D: IRB Approval

IRB Approval # 1612011748

Inbox x

**Taylor, Amy L.** <altaylor@towson.edu>

12/14/16 ★



to me, Saradha, IRB ▾

The IRB has approved your protocol "Frequency Following Response (FFR): Examining variability of this Complex Response " effective 12/14/2016

Your IRB protocol can now be viewed by your faculty advisor in MyOSPR. For more information, please visit: <http://www.towson.edu/academics/research/sponsored/myospr.html>

If you should encounter any new risks, reactions, or injuries to subjects while conducting your research, please notify IRB@towson.edu. 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.

We do offer training and orientation sessions for faculty/staff, please sign up for one of the sessions: <http://fusion.towson.edu/www/signupGeneric/index.cfm?type=OSPR>

Check back to that registration site frequently – we'll post additional sessions for January and spring semester soon.

Regards,
Towson IRB

References

- Ahadi, M., Pourbakht, A., Jafari, A. H., Shirjian, Z., & Jafarpisheh, A. S. (2014). Gender disparity in subcortical encoding of binaurally presented speech stimuli: An auditory evoked potentials study. *Auris Nasus Larynx*, 41, 239-243.
doi:10.1016/j.anl.2013.10.010
- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, 245, 35-47
doi:10.1016/j.heares.2008.08.004
- 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, 922-933.
doi:10.1016/j.clinph.2007.12.010
- Ananthakrishnan, S., Krishnan, A., & Bartlett, E. (2016). Human frequency following response: Neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear & Hearing (01960202)*, 37(2), e91-e103. doi:10.1097/AUD.0000000000000247
- Anderson, S., & Kraus, N. (2013). The potential role of the cABR in assessment and management of hearing impairment. *International Journal of Otolaryngology*, 2013, 1-10. doi:10.1155/2013/604729
- Batra, R., Kuwada, S., & Maher, V. L. (1986). The frequency-following response to continuous tones in humans. *Hearing research*, 21(2), 167-177.

- Belin, P. (2006). Voice processing in human and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1476), 2091-2107. doi:10.1098/rstb.2006.1933
- Bidelman, G. M. (2015). Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. *Hearing Research*, 323, 68-80. doi:10.1016/j.heares.2015.01.011
- Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, 1355, 112-125. doi:10.1016/j.brainres.2010.07.100
- Borden, G. J., Harris, K. S., & Raphael, L. J. (1994). *Speech science primer: Physiology, acoustics, and perception of speech*. Baltimore: Williams & Wilkins, c1994.
- Burkard, R., & Don, M. (2015). Introduction to auditory evoked potentials. In J. Katz, M. Chasin, K. English, L. Hood, & K. Tillery (Eds.), *Handbook of Clinical Audiology* (7th ed.) (pp. 187-205). Philadelphia: Wolters Kluwer Health.
- Cariani, P. (1999). Temporal coding of periodicity pitch in the auditory system: An overview. *Neural Plasticity*, 6(4), 147-172. doi:10.1155/NP.1999.147
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, 47(2), 236-246.
- Clark, W. W., & Ohlemiller, K. K. (2008). *Anatomy and physiology of hearing for audiologists*. Clifton Park, NY: Thomson Delmar, c2008.

- Clinard, C., Tremblay, K., & Krishnan, A. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hearing Research, 264*(1-2), 48-55.
doi:10.1016/j.heares.2009.11.010
- Davis, H., & Hirsh, S. K. (1976). The audiometric utility of brain stem responses to low-frequency sounds. *Journal of Auditory Communication, 15*(3), 181.
doi:10.3109/00206097609071775
- Emanuel, D. & Letowski, T. (2009). *Hearing Science*. Baltimore, MD: Lippencott, Williams & Wilkins.
- Fowler, C. G., & Horn, J. H. (2012). Frequency dependence of binaural interaction in the auditory brainstem and middle latency responses. *American Journal of Audiology, 21*(2), 190-198 9p. doi:1059-0889(2012/12-0006
- Galbraith, G., & Doan, B. (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.
doi:10.1016/0167-8760(95)00008-G
- Gardi, J., Merzenich, M., & Mckean, C. (1979). Origins of the scalp-recorded frequency-following response in the cat. *International Journal of Audiology, 18*(5), 353-380.
doi:10.3109/00206097909070062
- Gong, Q., Xu, Q., & Sun, W. (2013). Design and implementation of frequency-following response recording system. *International Journal of Audiology, 52*(12), 824-831.
doi:10.3109/14992027.2013.834537

- Greenberg, S., Marsh, J., Brown, W., & Smith, J. (1987). Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hearing Research, 25*(2-3), 91-114. doi:10.1016/0378-5955(87)90083-9
- Hoormann, J., Falkenstein, M., Hohnsbein, J., & Blanke, L. (1992). The human frequency-following response (FFR): Normal variability and relation to the click-evoked brainstem response. *Hearing research, 59*(2), 179-188.
- Jerger, J., & Hall, J. (1980). Effects of age and sex on auditory brainstem response. *Archives of Otolaryngology, 106*(7), 387-391.
- Jewett, D. L., & Williston, J. S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain: A Journal of Neurology, 94*(4), 681-696.
- Kiylioglu, N., Parlaz, A. U., Akyildiz, U. O., & Tataroglu, C. (2015). Evoked potentials and disability in multiple sclerosis: A different perspective to a neglected method. *Clinical Neurology and Neurosurgery, 133*, 11-17.
- Krishnan, A. (2002). Human frequency-following responses: Representation of steady-state synthetic vowels. *Hearing Research, 166*, 192-201. doi:10.1016/S0378-5955(02)00327-1
- Krishnan, A., & Gandour, J. T. (2014). Language experience shapes processing of pitch relevant information in the human brainstem and auditory cortex: Electrophysiological evidence. *Acoustics Australia, 42*(3), 166-178.
- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2012). Experience-dependent plasticity in pitch encoding: From brainstem to auditory cortex. *Neuroreport, 23*(8), 498-502. doi:10.1097/WNR.0b013e328353764d

- 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. doi:10.1016/j.cogbrainres.2005.05.004
- Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012a). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the United States of America*, 109(20), 7877-7881.
- Krizman, J., Skoe, E., & Kraus, N. (2010). Stimulus rate and subcortical auditory processing of speech. *Audiology and Neurotology*, 15(5), 332-342. doi:10.1159/000289572
- Krizman, J., Skoe, E., & Kraus, N. (2012b). Sex differences in auditory subcortical function. *Clinical Neurophysiology*, 123(3), 590-597.
- Kumar, K., Bhat, J. S., D'Costa, P. E., Srivastava, M., & Kalaiyah, M. K. (2013). Effect of stimulus polarity on speech evoked auditory brainstem response. *Audiology Research*, 3(1), 52-56. doi:10.4081/audiores.2013.e8
- Lehmann, A., & Schönwiesner, M. (2014). Selective attention modulates human auditory brainstem responses: Relative contributions of frequency and spatial cues. *Plos ONE*, 9(1), 1-10. doi:10.1371/journal.pone.0085442
- McFadden, D. (1998). Sex differences in the auditory system. *Developmental Neuropsychology*, 14(2-3), 261-298.

- Moore, B. J. (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.
- Moore, B., & Glasberg, B. (1989). Mechanisms underlying the frequency discrimination of pulsed tones and the detection of frequency modulation. *Journal of The Acoustical Society of America*, *86*(5), 1722-1732. doi:10.1121/1.398603
- Moushegian, G., Rupert, A., & Stillman, R. (1973). Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalography and Clinical Neurophysiology*, *35*(6), 665-667. doi:10.1016/0013-4694(73)90223-X
- Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *Plos ONE*, *9*(2), 1-23. doi:10.1371/journal.pone.0089642
- Munivrana, B., & Mildner, V. (2013). Cortical auditory evoked potentials in unsuccessful cochlear implant users. *Clinical Linguistics & Phonetics*, *27*(6/7), 472-483. doi:10.3109/02699206.2013.771214
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(40), 15894-15898. doi:10.1073/pnas.0701498104
- Nan, Y., Skoe, E., Nicol, T., & Kraus, N. (2015). Auditory brainstem's sensitivity to human voices. *International Journal of Psychophysiology*, *95*, 333-337. doi:10.1016/j.ijpsycho.2014.12.013

- Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging*, *33*(7), e1-e4. doi:10.1016/j.neurobiolaging.2011.12.015
- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience*, *29*(45), 14100-14107. doi:10.1523/JNEUROSCI.3256-09.2009
- Passmore, S. R., Murphy, B., & Lee, T. D. (2014). The origin, and application of somatosensory evoked potentials as a neurophysiological technique to investigate neuroplasticity. *Journal of The Canadian Chiropractic Association*, *58*(2), 170-183.
- Paulraj, M., Subramaniam, K., Yaccob, S., Bin Adom, A., & Hema, C. (2015). Auditory evoked potential response and hearing loss: A review. *Open Biomedical Engineering Journal*, *9*, 17-24.
- Picton, T. W. (2011). *Human auditory evoked potentials*. San Diego, California; Abingdon, England: Plural Publishing, Inc., 2011.
- Plourde, G. (2006). Auditory evoked potentials. *Best Practice & Research Clinical Anaesthesiology*, *20*, 129-139. doi:10.1016/j.bpa.2005.07.012
- Rose, J. E., Brugee, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, *30*(4), 769-793.

- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(37), 15516-15521.
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, *156*(1), 95-103. doi:10.1016/j.bbr.2004.05.012
- Russo, N. M., Skoe, E., Nicol, T., Zecker, S., Kraus, N., Trommer, B., & Bradlow, A. (2008). Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders. *Clinical Neurophysiology*, *119*(8), 1720-1731. doi:10.1016/j.clinph.2008.01.108
- Sayles, M., & Winter, I. (2008). Reverberation challenges the temporal representation of the pitch of complex sounds. *Neuron*, *58*(5), 789-801. doi:10.1016/j.neuron.2008.03.029
- Sęk, A., & Bukała, M. (2013). The use of the phase locking information in the human auditory system for frequencies above 5 kHz. *Acta Physica Polonica*, *123*(6), 1106-1113. doi:10.12693/APhysPolA.123.1106
- Skoe, E., Burakiewicz, E., Figueiredo, M., & Hardin, M. (2017). Basic neural processing of sound in adults is influenced by bilingual experience. *Neuroscience*. doi:10.1016/j.neuroscience.2017.02.049
- Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial. *Ear & Hearing*, *31*(3), 302-324. doi:10.1097/AUD.0b013e3181cdb272

- Smith, J., Marsh, J., & Brown, W. (1975). Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*, 39(5), 465-472. doi:10.1016/0013-4694(75)90047-4
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex*, 22(5), 1180-1190. doi:10.1093/cercor/bhr196
- Stapells, D. (2009). Cortical Event-Related Potentials to Auditory Stimuli. In J. Katz, L. Medwetsky, R. Burkard, and L. Hood (Eds.), *Handbook of Clinical Audiology* (6th Ed.) (pp. 395-415).
- Stegeman, D., Roeleveld, K., Dumitru, D., & King, J. (1997). Near- and far-fields: Source characteristics and the conducting medium in neurophysiology. *Journal of Clinical Neurophysiology*, 14(5), 429-442. doi:10.1097/00004691-199709000-00009
- Stillman, R., Crow, G., & Moushegian, G. (1978). Components of the frequency-following potential in man. *Electroencephalography and Clinical Neurophysiology*, 44(4), 438-446. doi:10.1016/0013-4694(78)90028-7
- Tremblay, K., Clinard, C. (2015). Cortical Auditory-Evoked Potentials. In J. Katz, M. Chasin, K. English, L. Hood, & K. Tillery (Eds.), *Handbook of Clinical Audiology* (7th ed.) (pp. 337-356). Philadelphia: Wolters Kluwer Health.
- Vander Werff, K. R., & Burns, K. S. (2011). Brain stem responses to speech in younger and older adults. *Ear & Hearing* (01960202), 32(2), 168-180. doi:10.1097/AUD.0b013e3181f534b5

- Walker, K. M., Bizley, J. K., King, A. J., & Schnupp, J. H. (2011). Cortical encoding of pitch: Recent results and open questions. *Hearing Research, 271*(1-2), 74-87.
doi:10.1016/j.heares.2010.04.015
- Wever, E. G., & Bray, C. W. (1930). Action currents in the auditory nerve in response to acoustical stimulation. *Proceedings of the National Academy of Sciences of the United States of America, 5*, 344-350.
- Worden, F. G., & Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography & Clinical Neurophysiology, 25*(1), 42-52. doi:10.1016/0013-4694(68)90085-0

CURRICULIM VIATE

NAME: Lindsay Roberts

PROGRAM OF STUDY: Audiology

DEGREE AND DATE TO BE CONFERRED: Doctor of Audiology (Au.D.), May, 2018

SECONDARY EDUCATION: Chesapeake Sr. High School

<u>Collegiate institutions attended</u>	<u>Dates</u>	<u>Degree/Date of Degree</u>
University of Maryland –College Park, MD Major: Hearing and Speech Sciences	8/2009 - 5/2013	B.A. 5/2013
Towson University—Towson, MD	8/2014 - 5/2018	Au.D. Anticipated 5/2018

PROFESSIONAL POSITIONS HELD:

**All position listed below were clinical internships in accordance with Towson University’s Audiology Doctoral Program

ENTAA Care January 2017 — May 2017
 Doctoral Student Clinician
 129 Lubrano Dr L100,
 Annapolis, MD 21401

Children’s Hospital of Philadelphia August 2016 – December 2016
 Doctoral Student Clinician
 3401 Civic Center Blvd.
 Philadelphia, PA 19104

Chesapeake Hearing Centers May 2016 - August 2016
 Doctoral Student Clinician
 580 Ritchie Hwy,
 Severna Park, MD 21146

Hearing and Speech Agency January 2016—May 2016
 Doctoral Student Clinician
 1813 York Road
 Lutherville-Timonium, MD 21093

Towson University Speech, Language, and Hearing Center August 2015- December 2015
 Doctoral Student Clinician
 One Olympic Place
 Towson, MD 21204

