Spring 2013

The effects of musical training on perception and neural representation of temporal fine structure

Verleyne Andrews-Rodgers
James Madison University

The Effects of Musical Training on Perception and Neural Representation of Temporal Fine Structure

Verleyne Andrews-Rodgers

A dissertation submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Doctor of Audiology

Communication Sciences and Disorders

May 2013

Dedication

This is dedicated to my husband Dr. Claudius Rodgers, (ABD), my parents Drs. Vernon and Phyllis Andrews, and my siblings, Vernetta, Vaughn, and Veldon (Donnie). Claudius, I think you are aware that words cannot describe how invaluable you were in being able to make this dream of doctoral education a reality. To my parents, who set the example and provided the inspiration for pursuing higher education, I am not sure if you are aware, but because of you I have always wanted to be called "Dr. Andrews". To my siblings...I share this with you because I know that you are proud of your baby sister's accomplishments, and I am equally proud of you guys!

Acknowledgments

I am grateful to so many individuals who, in different ways, contributed to this project. First of all, I must thank Dr. Clinard for his tremendous work, resources and direction with this study. His mentorship as an advisor is truly unmatched, and I appreciate the patience, positivity and professionalism in the way in which he provided feedback and guidance. My committee members, Dr. Lincoln Gray and Dr. Brenda Ryals were excellent in their support and in providing critical feedback. Thank you to Sophia Patrikis, my undergraduate assistant, who helped during the arduous hours of data collection, and to Caitlin Cotter who tediously organized my files and scanned my lab notebook. I would like to acknowledge the Roger Ruth Memorial Student Research Grant for providing funds for supplies. I must also make mention of the participants in this study who devoted their time to this project.

Of course, a huge thank you to my husband Claudius who was amazing in shouldering so much responsibility while I worked on getting this done. His support and encouragement made it possible to keep going. To all of my family and friends who assured me many times that they were praying and sending well wishes my way. This truly helped to keep me sane throughout the process.

Last but not least, it would be remiss of me not to acknowledge the Almighty God, who provided strength and motivation to reach to this point in my academic and professional career. I truly believe that it was His plan to bring me to JMU and He provided what was necessary in order to complete this journey successfully.

Table of Contents

Dedication ii

Acknowledgments iii

Table of Contents iv

List of Tables vi

List of Figures vii

Abstract viii

I. Introduction 1

II. Review of Literature 2

- Introduction. 2
- Behavioral TFS processing is related to speech-in-noise understanding. 3
- Perceptual effects of musical training. 6
  - Musicians have been associated with improved auditory perception. 6
  - Behavioral TFS processing can be improved with training. 6
  - The effects of musical training on speech-in-noise understanding. 9
- Physiological effects of musical training. 9
- Conclusion. 13

III. Materials and Methods 15

- Subjects. 15
- Stimuli. 16
- Procedure 17
  - Behavioral Frequency Discrimination. 17
  - Speech Perception in Noise. 19
  - Frequency-Following Responses 20
- Statistical Approach. 23

IV. Results 25

- Frequency Discrimination. 25
- Frequency-Following Response. 26
- Speech-in-Noise. 27
- Relationships between measures. 28
  - FDL x FFR correlations. 29
  - FDL x HINT correlations 31
  - FFR x HINT correlations. 31
- Years of musical training x measures. 34
- Resistance to noise. 36

V. Discussion 38
Musical experience on frequency discrimination. 38
Musical experience and the FFR. 39
Musical experience and Speech-in-Noise. 41
Brain-Behavior Relationships/Relationships between measures. 42
Methodological Issues. 44
Clinical Applications. 45
Future Directions. 46
Conclusions. 47
VI. References 48

List of Tables

Table 1. Correlations between Years of Musical Training and FDLs. 36

Table 2. Correlations between Years of Musical Training and FFR PC. 36

Table 3. Correlations between Years of Musical Training and FFR Amp 36

List of Figures

Figure 1. Example Output of a Fast Fourier Transform 17

Figure 2. FFR Data for an Individual Participant 22

Figure 3. FDLs Shown by Frequency, Group, and Noise Conditions 25

Figure 4. FFR Phase Coherence Shown by Frequency, Group, and Noise Conditions.... 26

Figure 5. Box Plot Showing HINT Signal-to-noise ratios (SNRs) 28

Figure 6. Scatter Plots between FDLs and FFR Phase Coherence 30

Figure 7. Scatter Plots between HINT SNRs and FDLs 32

Figure 8. Scatter Plots between HINT SNRs and FFR Phase Coherence. 33

Figure 9. Scatter Plots between Years of Musical Experience and FDLs 35

Figure 10. Box Plots Comparing FFRs between Published Data and Present Study 41

Abstract

One of the most common complaints of persons with sensorineural hearing loss is difficulty hearing in background noise. Temporal fine structure (TFS) is one of the factors that contributes to understanding speech in the presence of background noise. TFS refers to the periodic information in speech which helps us to identify which speech sound we are listening to. TFS is also negatively affected by hearing loss, as well as age. In a quest to discover how TFS processing and thus speech-in-noise understanding can be improved, this study examined the effects of musical training on behavioral and physiological measures of temporal fine structure, as well as the brain-behavior relationship as it relates to frequency representation in the brainstem.

This relationship was measured by two behavioral tests: frequency discrimination and a measure of speech understanding in background noise – the Hearing-in-Noise test (HINT), and one physiologic measure, the frequency following response (FFR). The stimuli for frequency discrimination and the FFR were tonebursts of 500 Hz in quiet, 1000 Hz in quiet, 500 Hz in noise, and 1000 Hz in noise. A total of 28 subjects were tested, 16 musicians and 12 non-musicians.

The results showed that musicians had better frequency difference limens (FDLs) than non-musicians. For the physiologic measure, musical experience did not affect phase-locked representations of TFS. Musicians also did not have better signal-to-noise ratios on the HINT. There were no significant brain-behavior relationships between measures except that lower or better FDL thresholds at 1000 Hz in quiet implied lower or worse phase coherence at 1000 Hz in quiet. A greater number of years of musical experience related to lower or better FDLs for the conditions in quiet but not in noise.

The years of training did not relate to performance on FFR phase coherence, amplitude, or HINT scores. It was concluded that musical training significantly enhanced behavioral TFS processing, however no significant effects were noted for neural representation of TFS or speech-in-noise understanding.

Chapter I

Introduction

The purpose of this study was to examine the effects of musical training on behavioral and physiological measures of temporal fine structure, as well as to examine the brain-behavior relationship as it relates to frequency representation in the brainstem. This was measured by two behavioral tests, frequency discrimination and a measure of speech understanding in background noise – the Hearing-in-Noise test (HINT), and one physiologic measure, the frequency following response (FFR). It was hypothesized that musicians would demonstrate smaller (better) frequency discrimination difference limens (FDL) than non-musicians in quiet and in noise. It was also hypothesized that musicians would have lower (better) HINT signal-to-noise ratios (SNRs) than non-musicians; and that musicians would demonstrate stronger phase coherence and amplitude FFR measures than non-musicians in quiet and in noise.

Chapter II

Review of Literature

Introduction.

Humans can improve their performance on many perceptual tasks with practice. This learning demonstrates that our perceptual systems are not rigid, but in fact can be modified through short- or long-term listening experience (e.g., Wright, 2001). This can be seen over the course of musical training, where one can see an improvement of auditory skills as it relates to the perception of frequency, timing, intensity, timbre, etc. (Draganova, Wollbrink, Schulz, Okamoto, & Pantev, 2009; Geiser, Sandmann, Jäncke, & Meyer, 2010; Kraus, Skoe, Parbery-Clark, & Ashley, 2009; Geringer, 1995). It is unclear however, whether improvement through training is related to changes in cortical and/or brainstem processing of sounds.

Speech has complex acoustic attributes and it is composed of various acoustic events, such as changes in frequency, intensity, and timing cues. This dissertation looks specifically at effects of musical training on behavioral and physiological representation of frequency. Rosen (1992) proposed a model of describing the temporal information in speech, using envelope, periodicity, and fine structure. Envelope information can be defined as the relatively slow fluctuations in overall amplitude from 2-50 Hz. In speech, the envelope gives segmental cues to the manner of articulation, voicing, and vowel identity, as well as prosodic cues (Rosen, 1992). Periodicity of a speech signal relates to the distinction between periodic and aperiodic stimulation, as well as the rate of periodic stimulation. Periodic fluctuations primarily range from 50-500 Hz. Fine structure, or temporal fine structure (TFS), refers to the variations of wave shape within single periods of periodic sounds, or over short time intervals of aperiodic sounds. According to Rosen (1992) fine structure information in speech consists of energy above approximately 500 Hz. TFS gives acoustic information about the spectrum (amplitude and phase) of a sound, as well as its formant pattern, which correlates to timbre and quality.

Behavioral TFS processing is related to speech-in-noise understanding.

Speech-in-noise understanding is a complex task which is related to different factors which include but are not limited to peripheral auditory sensitivity, attention, and temporal fine structure processing (Tremblay, Picard, Barbarosie, & Banville, 1991; Humes, 1996; Buss, Hall, & Grose, 2004; Strelcyk & Dau, 2009). Of particular interest, in this document, is temporal fine structure (TFS), which has been investigated in order to determine its importance in speech recognition. One way that behavioral TFS processing has been related to speech perception is by using psychophysical tasks with stimuli designed to focus the listener on TFS, rather than envelope information, and comparing those results to speech-in-noise measures. Some examples of these behavioral TFS tests include frequency discrimination (Abel, Krever, & Alberti, 1990; Micheyl, Moore, & Carlyon, 1998; Clinard, Tremblay, & Krishnan, 2010), frequency modulation detection limens (Buss et al., 2004), and detection of interaural phase differences (Grose & Mamo, 2010; Hopkins & Moore, 2011; Sęk & Moore, 2012). Another way that behavioral TFS contributions to speech perception have been studied involves manipulating the TFS of speech, through processes such as vocoding, and examining how altered TFS affects speech understanding (Apoux & Healy, 2011; Drennan, Won, Dasika, & Rubinstein, 2007). Taken together, these studies indicate that behavioral measures of TFS processing are closely related to speech understanding in noise.

In looking at the effect of age on behavioral TFS processing, Hopkins and Moore (2011) as well as Neher, Lunner, Hopkins, and Moore (2012) reported that older, normal-hearing listeners showed significantly poorer performance on tasks measuring TFS sensitivity than younger normal hearing listeners. They found that behavioral interaural phase difference (IPD) thresholds were significantly correlated with speech understanding in background noise (particularly when there are modulations in the noise amplitude), such that better IPD thresholds indicated better the speech-in-noise understanding.

Apoux and Healy (2011) were interested in the relative contributions of the TFS in the target speech stimuli and the TFS of a masker (multi-talker babble). By manipulating the TFS of both target and masker simultaneously and independently using a vocoder, they were able to determine that listeners rely primarily on the target-speech TFS, and that the masker TFS has very little influence on speech understanding (Apoux & Healy, 2011).

Drennan et al. (2007) manipulated the percentage of TFS randomization in speech to demonstrate its importance when listening to speech in noise. Using cochlear implant simulations in normal-hearing listeners, they studied the contribution of TFS to speech understanding in noise, by including varying degrees of intact and vocoder-randomized TFS in the speech stimuli. They found that speech recognition thresholds in noise were best with intact TFS and that these thresholds became poorer as the percentage of randomized TFS increased. Randomization of TFS significantly reduced speech understanding in the presence of speech-shaped noise and multi-talker babble.

It is also no surprise that individuals with sensorineural hearing loss struggle with speech understanding in the presence of background noise. Thus, with the discovery of the effect of TFS on speech recognition in noise, the question of the effects of hearing loss on TFS processing was a natural progression. Buss et al. (2004) compared normal-hearing listeners to those with sensorineural hearing loss. They found that the participants with hearing loss had poorer performance in both speech recognition and psychophysical tasks, and that there were significant correlations within the hearing loss group when looking at speech recognition performance and a psychophysical task measuring frequency modulation detection.

Similarly, Strelcyk and Dau (2009) also found that hearing impaired listeners had poorer performance on frequency selectivity, TFS processing, and speech reception than normal hearing listeners. This paper showed significant correlations between behavioral TFS processing and speech-in-noise understanding.

This section makes the point that behavioral TFS processing is related to speech-in-noise understanding. It was also shown that there are populations that have difficulty understanding speech-in-noise, such as individuals with sensorineural hearing loss. If there were a way to improve behavioral TFS processing, then maybe there’s a way to improve speech-in-noise perception in those populations. For several years, musical training has been touted as a tool which has overall and specific benefits on auditory perception. As such, the aim of this study is to examine whether there are connections between musical training, TFS processing and speech-in-noise understanding.

Perceptual effects of musical training.

Musicians have been associated with improved auditory perception.

Musicians have been reported to have perceptual advantages over non-musicians on a variety of listening tasks. Research has shown that musicians have a listening advantage on tasks such as frequency discrimination (Kishon-Rabin, Amir, Vexler, & Zaltz, 2001; Micheyl, Delhommeau, Perrot, & Oxenham, 2006; Nikjeh, Lister, & Frisch, 2008; Akin & Belgin, 2009), gap detection (Wright, 2001; Zendel & Alain, 2011), detection of intensity changes (Geringer, 1995), perception of timbre (Seither-Preisler et al., 2007), and tasks designed to test auditory memory (Pallesen et al., 2010; Tierney, Bergeson, & Pisoni, 2008). In addition, musicians have been reported to have better speech understanding than non-musicians in the presence of background noise (Parbery-Clark, Strait, Anderson, Hittner, & Kraus, 2011; Zendel & Alain, 2011).

Behavioral TFS processing can be improved with training.

Perception of TFS is malleable. This has been demonstrated through studies that have used psychophysical training as well as studies that have examined musically-trained individuals. Taken together, these studies demonstrate that the perception of TFS can be improved by multiple types of training paradigms. Although the following paragraphs will provide studies that show how the perception of TFS can be improved, the physiological basis for this improved perception is not completely understood. Frequency discrimination tasks are used in order to assess behavioral TFS processing, and a number of studies have reported that training, musical or psychophysical, improves frequency discrimination (Campbell & Small, 1963; Spiegel & Watson, 1984; Demany, 1985; Irvine, Martin, Klimkeit, & Smith, 2000; Wright & Ortiz, 2000; Kishon-Rabin et al., 2001; Micheyl et al., 2006; Nikjeh et al., 2008; Akin & Belgin, 2009; Carcagno & Plack, 2011).

An early study by Spiegel and Watson (1984) compared the auditory frequency discrimination abilities of professional musicians with those of non-musicians. The stimuli consisted of single tones or tone patterns. Using single tones, only half of the non-musicians had difference thresholds as low as the musicians, whereas the other half of non-musicians had higher or worse thresholds than the musicians. For the tone pattern stimuli, the musicians' median difference thresholds were about three times smaller than non-musicians.

Kishon-Rabin et al. (2001) were interested in the effect of musical training on frequency discrimination. As such, they compared a group of professional musicians to non-musicians by using adaptive forced-choice procedures to determine FDLs. They found that musicians had significantly better difference limens than non-musicians at 250, 1000, and 1500 Hz. The effect of years of musical training was significantly correlated to FDL in the musicians, whereby a greater number of years correlated to better FDL.

In a study looking at pitch discrimination of musicians and non-musicians, Nikjeh and colleagues (2008) used a psychoacoustic component along with an electrophysiologic measure mentioned later. They examined FDLs between musicians and non-musicians and found a significant difference between groups, where the FDLs for musicians were 50% smaller than those of the non-musicians.

Akin and Belgin (2009) conducted a study comparing FDLs as well as frequency modulation difference limens (FMDLs) of musicians and non-musicians. The study, using 16 musicians and 16 non-musicians, yielded the conclusion that musicians had significantly better frequency discrimination on both the FDL and FMDL tasks at frequencies in the range of 125-8000 Hz.

In a similar study, Micheyl et al. (2006) looked at the influence of musical and psychophysical training on pitch discrimination. In their first pitch discrimination task, they found that non-musicians had thresholds six times poorer than musicians and only four times poorer after two hours of psychophysical training. In a second experiment, the investigators trained eight non-musicians (not part of the first experiment) for 14 hours and found that it took four to eight hours of training for their pitch discrimination thresholds to be as small as the classical musicians in the first experiment. Their findings support the idea that frequency discrimination can be improved with training, whether that training was related to classical musical training or traditional psychophysical training.

The effect of perceptual learning on frequency discrimination was studied by Demany (1985), by using a frequency discrimination psychophysical task. Frequency discrimination thresholds were determined for a 200 Hz tone before and after a psychophysical training task. For the training, participants were divided into four groups, each trained on a different frequency (200, 360, 2500 or 6000 Hz). After training, the results showed a trend where frequency discrimination was improved for all of the training frequencies, with 200, 360 & 2500 Hz being significantly better than 6000 Hz at improving frequency discrimination at 200 Hz.

Previously, it was shown that several studies have identified a connection between temporal fine structure and speech in noise understanding. The above studies also demonstrate that musicians have superior frequency discrimination abilities to non-musicians. As such, the question of whether musicians also display better speech-in-noise processing will be discussed below.

The effects of musical training on speech-in-noise understanding.

Several recent studies have examined speech-in-noise understanding of musicians compared to non-musicians. If musical training improves speech understanding in challenging listening environments, then aural rehabilitation of populations with difficulty in those environments (e.g., older adults with or without hearing loss) may potentially benefit from a better understanding of how musical training and speech perception interact. Parbery-Clark and colleagues have conducted several studies on the effect of musical training, particularly as it relates to speech-in-noise understanding (Parbery-Clark, Skoe, & Kraus, 2009; Parbery-Clark, Skoe, Lam, & Kraus, 2009; 2011b). Their studies report that musicians have enhanced speech-in-noise understanding as compared to non-musicians. Additionally, a comparison of speech-in-noise understanding in younger and older musicians and non-musicians supported this finding (Parbery-Clark et al., 2011b).

Physiological effects of musical training.

In an attempt to better understand the effects of musical training on the auditory system, researchers have not only used behavioral tests, but physiological measures as well. This section reviews studies that have examined effects of musical training on electrophysiological measures of frequency representation. Some of these studies report that musicians have enhanced auditory evoked potentials when compared to non-musicians. These differences are seen in studies primarily using cortical auditory evoked potentials (e.g., mismatch negativity) as well as several studies that used brainstem auditory evoked potentials (i.e., frequency-following response). It should be noted that although stronger phase-locked representations of the temporal envelope have been reported in musicians, neural representation of TFS in musicians has not yet been addressed by the literature.

A number of studies have used cortical auditory evoked potentials, such as the mismatch negativity or P300 response, to demonstrate group differences between musicians and non-musicians. These two cortical responses indicate that, at the level of the auditory cortex, the responses of musicians are more sensitive to fine differences in stimulus parameters such as frequency (Koelsch, Schröger, & Tervaniemi, 1999; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005; Herholz, Lappe, & Pantev, 2009; Wang, Staffaroni, Reid, Steinschneider, & Sussman, 2009; Trainor, Desjardins, & Rockel, 2011), duration (Tervaniemi et al., 2009; Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009), and intensity (Tervaniemi et al., 2009) in the context of an oddball stimulus paradigm. Although these responses indicate that the central auditory nervous systems of musicians have a greater capacity for detecting small contrasts between stimuli, these cortical responses do not indicate how well the stimulus is encoded. It is currently unknown whether these group-related differences in cortical responses are related to differences in sound encoding in the auditory brainstem.

The frequency-following response (FFR) is the product of a continued neural response to periodic stimuli, whereby the response is phase-locked to the individual frequencies within the stimulus waveform and/or the envelope of the stimuli (Krishnan, 2007). The FFR can actually give information about how acoustic aspects of stimuli (e.g., frequency-specific) are represented in the central auditory nervous system (CANS). The response originates in the brainstem and has multiple neural generators including the cochlear nuclei (CN), medial superior olive (MSO), and the inferior colliculus (IC). When recorded with a vertical electrode array, the primary neural generator of the FFR is the inferior colliculus (Smith, Marsh, & Brown, 1975). In contrast to the cortical responses described above, the FFR can reflect the quality of stimulus encoding for the temporal envelope and/or the fine structure. In addition, the FFR can also be used as a tool to examine brainstem plasticity at the level of the auditory brainstem.

Short-term listening experience on a speech-identification training procedure has been reported to enhance phase-locked stimulus representation as reflected by the FFR (Song, Skoe, Wong, & Kraus, 2008). Participants underwent two weeks of identification training on English syllables with mandarin tone contours applied to them. Post-training FFRs revealed an improvement in the ability of the FFRs to encode the pitch contour of the stimuli, or fundamental frequency.

The brainstem's capacity to improve phase-locked representations of stimuli has also been reported in two populations that have long-term listening experience: native tonal-language speakers (Krishnan, Gandour, Bidelman, & Swaminathan, 2009; Krishnan, Gandour, & Bidelman, 2010a, 2010b; Bidelman, Gandour, & Krishnan, 2011) and musicians (Wong, Skoe, Russo, Dees, & Kraus, 2007; Musacchia, Sams, Skoe, & Kraus, 2007; Musacchia, Strait, & Kraus, 2008; Lee, Skoe, Kraus, & Ashley, 2009; Parbery-Clark et al., 2009a, 2011b). Investigators have looked at whether native language experience has an effect on music-related processing and its perception. Bidelman et al. (2011) compared FFRs of native English-speaking musicians, native

English-speaking non-musicians, and native speakers of Mandarin Chinese. They used Western musical chords, and detuned musical chords (where the third of the chord was either sharp or flat) as stimuli for the FFR, and found that relative to non-musicians, both musicians and native Mandarin speakers had stronger brainstem representation of the fundamental frequency. Additionally, the studies that have reported FFRs in musicians have focused on the temporal envelope as well as how robust the response is to the presence of noise and have reported that musicians have enhanced FFRs to the temporal envelope when compared to non-musicians.

Thus far, one study has addressed the question of whether these music-training related FFR enhancements in quiet also apply to stimuli that are in the presence of noise. This study found that the temporal envelope representation in musicians is more robust to masking noise than non-musicians, providing the first and only comparison between musicians' perceptual advantage in speech-in-noise and the FFR. Parbery-Clark et al. (2009a) investigated this effect by comparing FFRs to a synthetic /da/ consonant vowel stimulus, sometimes referred to as complex auditory brainstem responses (cABRs) or speech-evoked ABRs, in quiet and in noise between musicians and non-musicians. The musicians were found to have more robust subcortical representations of the acoustic stimulus in the presence of noise than non-musicians as demonstrated by earlier onset response timing, as well as greater phase-locking amplitude to the temporal envelope, or fundamental frequency, of the waveform. In other words, the musicians had significantly less degraded response morphology in noise than the non-musicians. These neural measures were also associated with better behavioral performance on the Hearing in Noise Test (HINT), where the musicians demonstrated better scores than non-musicians.

As seen in the above studies, the literature has focused on the neural encoding of the temporal envelope in musicians, however not much is known about the neural TFS processing in musicians. This present study seeks to shed light in this area.

Conclusion.

In summary, it has been shown that TFS is a valuable contributor to speech-in-noise understanding, and that TFS can be improved by training. Studies have also shown that musicians have enhanced auditory processing, both behaviorally and physiologically, and display superior behavioral TFS processing than non-musicians. It was also reported by Parbery-Clark and colleagues (Parbery-Clark et al., 2009a; 2009b; 2011b) that musicians have enhanced speech-in-noise understanding over non-musicians. The question of whether brainstem encoding of TFS improves with musical training still remains however.

This dissertation study looks at the effects of musical training on behavioral and physiological measures of temporal fine structure as well as the brain-behavior relationship as it relates to frequency representation in the brainstem. The following hypotheses were addressed in this study:

(1) It was hypothesized that musicians would have lower (better) FDLs than non-musicians in quiet and in noise, indicating that musicians were superior at behavioral frequency discrimination.

(2) It was expected that musicians, compared to non-musicians, would be better able at understanding speech in the presence of background noise, as reflected by the musicians having lower (better) HINT SNRs than non-musicians.

(3) It was expected that the musicians, when compared to the non-musicians, would have better FFR responses in quiet and in noise, as demonstrated by stronger phase coherence and amplitude FFR measures.

(4) It was also hypothesized that there would be significant relationships between behavioral and physiological measures. Lower (better) FDLs were expected to be associated with higher (better) FFR phase coherence, and vice versa. Lower (better) HINT SNRs were also expected to be associated with higher (better) FFR phase coherence.

Chapter III

Materials and Methods

Subjects.

Twenty-eight subjects (ages 21-35, mean age 22.25 years, standard deviation = 2.80; 8 male, 20 female) participated in this study. All subjects had clinically normal hearing sensitivity, defined as thresholds ≤ 25 dB HL at octave frequencies from 250 to 8000 Hz. Case histories, hearing screenings, and acoustic immittance testing revealed that all subjects were right-handed, native, monolingual English speakers, had normal tympanometric measures, had no history of otological or neurological disorders, and were not taking interfering prescription medications. Two participants were excluded; one female non-musician who had a perforated left tympanic membrane, and one male musician who reported taking Attention Deficit-Hyperactivity Disorder medication.

The subjects were divided into two groups: musicians (n = 16, mean age 22.00 years, standard deviation = 2.31; 7 males, 9 females) and non-musicians (n = 12, mean age 22.58 years, standard deviation = 3.42; 1 male, 11 females). Musicians were defined as having at least ten years of musical training, either formal or informal, with self-reported consistent practice. Non-musicians were defined as those with less than three years of musical training, and any musical training had to have taken place at least seven years prior to enrollment in the study. Definitions of these groups are consistent with the published literature (Micheyl, Delhommeau, Perrot, Oxenham, 2006; Parbery-Clark, Skoe, Kraus 2009a; 2009b).

Subjects were recruited from James Madison University's students and staff primarily through posted fliers, word-of-mouth, and a class presentation. Each subject was compensated $10 per hour for their participation. All procedures in this study were approved by the institutional review board at James Madison University.

Stimuli.

Stimuli for the behavioral frequency discrimination and physiological conditions were toneburstsof 300 ms duration, including 15 ms rise/fall time shaped by a Hanning window. The physiological and behavioral conditions included four stimuli: 500 Hz toneburst (500 Hz in quiet), 1000 Hz toneburst (1000 Hz in quiet), 500 Hz toneburst in noise (500 Hz in noise), and a 1000 Hz toneburst in noise (1000 Hz in noise). In the physiological condition, the tonal stimuli were presented at 80 dB SPL, with a 20 dB signal-to-noise ratio for the noise condition using a third octave-wide narrowband noise, and in the behavioral condition, the tonal stimuli were presented at 70 dB SPL, also with a 20 dB signal-to-noise ratio for the noise condition using an octave-wide narrowband noise centered on the test frequencies. Signal-to-noise ratios were calibrated using the spectrum level of the noise (Figure 1), rather than RMS dB SPL, consistent with the frequency discrimination in noise literature (Dye & Hafter, 1980; Hienz, Sachs, & Aleszczyk, 1992; Plack, Turgeon, Lancaster, Carlyon, & Gockel, 2011). Onset polarity was positive. Stimuli were generated with a sampling frequency of 44.1 kHz. A magnetically shielded ER3-A insert earphone delivered stimuli to the right ear for both the physiological and behavioral conditions.

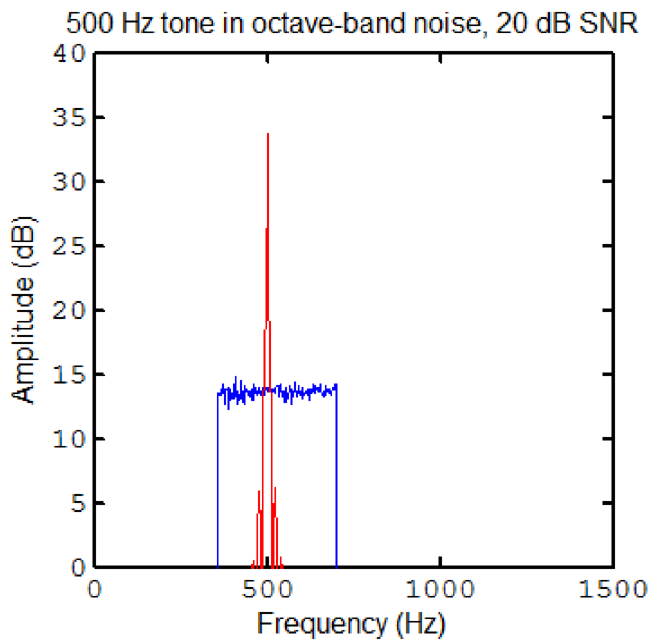

Figure 1. Example output of a fast Fourier transform illustrating how the signal-to-noise ratio was calibrated using Spectrum level (Lps), rather than RMS dB SPL. This calibration method follows the frequency-discrimination-in-noise methods in the literature (e.g., Dye and Hafter, 1980).

Procedure

Data collection consisted of three measures: one physiological measure - Frequency Following Response (FFR), and two behavioral measures – Frequency Difference Limen (FDL) test, and the Hearing-in-Noise Test (HINT). The test order and the order of conditions within each test were randomized. Data collection was typically performed in one session of approximately two and a half to four hours.

Behavioral Frequency Discrimination.

The frequency discrimination task was tested separately for each of the four stimuli (500 Hz in quiet, 500 Hz in noise, 1000 Hz in quiet, and 1000 Hz in noise), using an adaptive two-interval, two-alternative forced choice procedure with a two-down, one-up adaptive rule (Levitt, 1971). This procedure is based off of Olsho, Koch, and Carter (1988). A custom MATLAB (version 7.5; MathWorks, Natick, MA) program, used by Clinard et al., (2010), was developed for this procedure. During testing, all subjects were seated in a double-walled, sound-attenuating booth.

In each of the trials, a warning light with 300 ms duration preceded each pair of tones. Each pair of tones consisted of the frequency (e.g. 1000 Hz) and another tone which was always lower than the test frequency by a given amount, Δf. The order of these tones for each trial was randomized, with an inter-stimulus interval of 300 ms. Each subject was instructed to mouse-click on a button on the computer monitor which corresponded to the tone that had the higher pitch. After the subject's selection, the button that corresponded to the interval that had the higher pitch would briefly light up to indicate whether the correct interval was chosen or not. In the first trial, there was a large Δf of 5%, which resulted in 25 and 50 Hz for 500 and 1000 Hz, respectively. All subjects were able to discriminate this first pair. If the correct answer was chosen for two consecutive trials, the Δf decreased by half the previous value, whereas, if the incorrect answer was chosen, Δf doubled. This procedure continued until there were ten reversals. The frequency discrimination threshold was computed as the mean of the last eight reversals.

A minimum of two runs for each of the four stimuli (500 Hz in quiet, 500 Hz in noise, 1000 Hz in quiet, and 1000 Hz in noise) was collected. The mean frequency discrimination threshold for each of the stimuli was calculated and used to obtain the FDL which was the Δf/f.

Speech Perception in Noise.

The Hearing-in-Noise Test (HINT) was administered with both the speech and noise originating from a speaker in front of the listener at 0º azimuth; the center of the speaker was at ear-level for the listener of average height, when seated - approximately 39 inches above the floor. The test was presented by routing the output of a CD-player through a Grason-Stadler Instruments 61 (GSI-61) audiometer in a double-walled, sound-attenuating booth. The listener's head was positioned at a distance approximately 1.5 meters from the speaker where speech-shaped noise was at a calibrated level of 65 dB(A). The subject was instructed to listen to a male voice reading a sentence in the presence of background noise and repeat what he said, even if the man's voice seemed soft and difficult to understand, or even if the sentence did not make sense. The subject was also told that some sentences will be repeated and that they should continue to repeat the sentence.

Two phonemically balanced 10-sentence lists were randomly assigned by the tester, for a total of 20 sentences per participant in order to increase the number of reversals per adaptive track. The sentences were derived from the Bamford-Kowal-Bench sentences. In one channel, the noise was fixed at the dial setting 67 dB, which corresponded to 65 dB (A), and the speech stimuli were presented through the second channel. The first sentence was presented at 4 dB below the noise level (61 dB), and its level was increased in 4 dB steps until the sentence was repeated with 100% accuracy. After the final presentation level for the first sentence was recorded on the score sheet, the presentation level was decreased by 4 dB for the second sentence. If the response was correct (100% accuracy), then a "+" was marked on the score sheet for that sentence, and the presentation level was decreased by 4 dB. If the response was incorrect, then a "-" was marked for that sentence, and the presentation level of the sentence was increased by 4 dB. The 4 dB steps were used between sentences 1-3, and then 2 dB steps were used for sentences 4-20. Although 20-sentences were presented, the presentation level for a 21st sentence was included on the score sheet at the level it would be presented at were the test to continue.

The signal-to-noise ratio (SNR) was calculated for each participant. To calculate the SNR, the mean of presentation levels of sentences 5-21 was found (Reception Threshold of Speech-RTS), and then the noise level of 67 dB was subtracted from the mean in order to arrive at the SNR. The SNR corresponds to the 50% correct point.

Frequency-Following Responses

FFRs were collected via a Neuroscan Synamps 2 acquisition system using a single-channel recording (Krishnan, Xu, Gandour, & Cariani, 2005; Swaminathan, Krishnan, Gandour, & Xu, 2008). The subjects were in a double-walled, sound-attenuated booth, and were seated in a reclining chair. They were instructed to relax quietly. The responses were recorded from Cz (vertex) to the nape-of-the-neck, and the ground electrode was located on the left mastoid, with electrode impedances below 5 kΩ. Inter-electrode impedances were within 1 kΩ. Similar to the inter-stimulus interval (ISI) of the FDL task, the ISI was 300 ms for FFR conditions. The online electroencephalography (EEG) filters were 100–3000 Hz, the analysis time window was 0–320 ms, and the analog-to-digital sampling rate was 20 kHz. For each stimulus, one thousand individual artifact-free responses or sweeps were collected. Artifact rejection was set to reject any sweeps with a voltage of ±30 μV.

FFRs were collected separately to each of the four stimuli: 500 Hz in quiet, 500 Hz in noise with 20 dB SNR, 1000 Hz in quiet, and 1000 Hz in noise. Noise conditions had a 20 dB SNR. To minimize phase locking to the envelope of the noise, 100 unique noises were generated for each FFR-in-noise condition. These noises were combined with the same tone, then the 100 tone-in-noise stimuli were presented in a randomized order until the target number of accepted sweeps was met. A five-minute break was given between each FFR condition.

An offline analysis of the FFR data was conducted in two traditional ways: amplitude and phase coherence (PC) (Levi, Folsom, & Dobie, 1995). Custom MATLAB programs were developed for both amplitude and phase coherence analyses, which were performed on the output of a Fast Fourier Transform (FFT) and were used as statistical detection algorithms to verify response presence. In order to obtain FFT resolution of 0.96 Hz, consecutive pairs of sweeps were concatenated (John, Lins, Boucher, & Picton, 1998). The amplitude refers to the averaged magnitude of the neural response, whereas the PC indicates the degree of phase locking to the stimulus frequency.

In order to calculate the amplitude, the concatenated sweeps (500 double-sweeps) were averaged and submitted to an FFT (Figure 2). Using coherent sampling, each of the four stimulus frequencies was precisely specified, which limited the FFR to one bin, or point, in the FFT output. An integer number of cycles for each stimulus frequency was represented in the analysis window. If present, the FFR amplitude was obtained from the FFT bin where the averaged response was located. The mean of five FFT bins above and below the response bin was used to estimate the background noise, which corresponded to ±5 Hz. The signal-to-noise ratio resulting from the amplitude and noise measures, was used as an F-ratio with 2, 20 degrees of freedom (Dobie & Wilson, 1996). The response was noted to be present if the amplitude was determined to be significantly greater than the background noise using a p-value < 0.05.

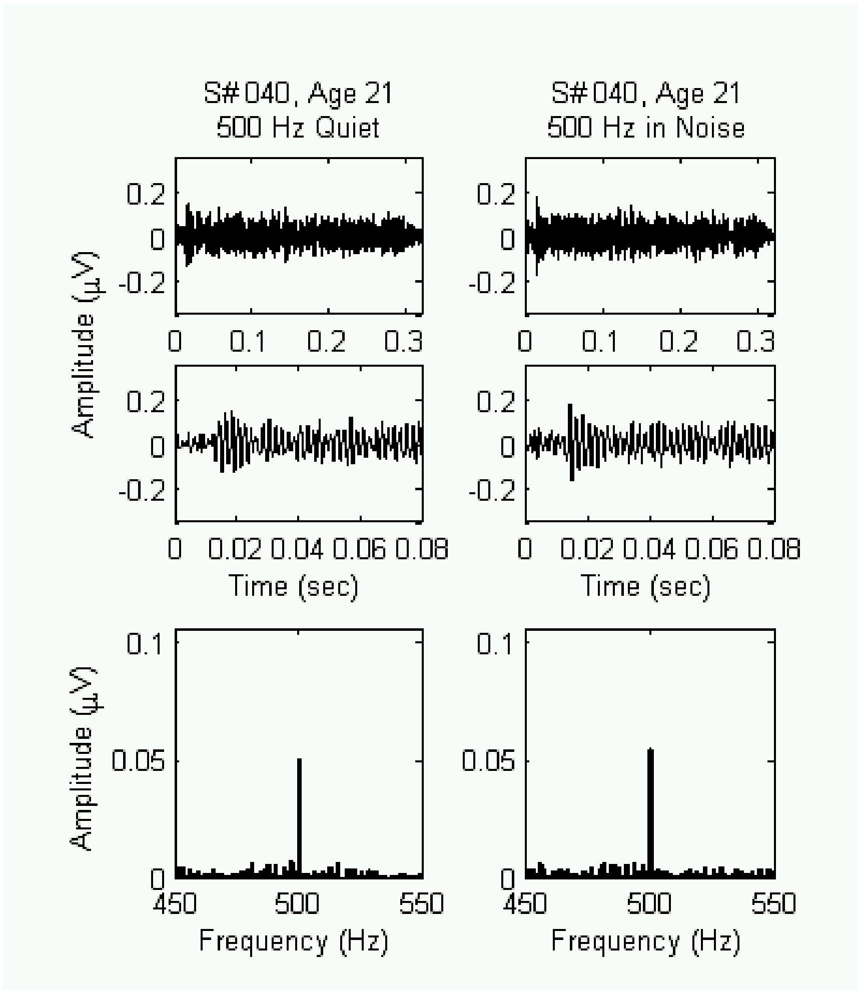

Figure 2. FFR data for an individual participant at 500 Hz in quiet (left panel) and 500 Hz in noise (right panel). Top rows show raw EEG waveforms. Bottom row shows FFT output of the FFR amplitude.

The non-averaged, concatenated sweeps were used to calculate the phase coherence via FFT analysis. Phase data from the stimulus frequency's FFT bin was compared across sweeps to then statistically determine the degree of phase locking to the stimulus frequency (Dobie & Wilson, 1989). The range of values for PC is from 0 to 1, with 0 being no phase locking or random phase across sweeps, and 1 being perfect phase locking. In order to assess circular uniformity the Rayleigh test (Fisher, 1993) was used, which indicated the response presence by a p-value < 0.05.

Statistical Approach.

Six analyses of variance (ANOVAs) were performed. Three-factor ANOVAs were conducted with either FDLs, FFR phase coherence, or FFR amplitude as the dependent variables; factors were Group (between-subjects on 2 levels, musicians and non-musicians), frequency (within-subjects on two levels, 500 and 1000 Hz), and noise (within-subjects on two levels, quiet and noise). The fourth ANOVA was done with HINT SNRs as the dependent variable, and Group as the factor (between-subjects on 2 levels, musicians and non-musicians). Partial η^2 was used for a measure of effect size, with small, medium, and large effect sizes defined as 0.0099, 0.0588, and 0.1379, respectively (Cohen, 1988).

In order to assess whether the FFRs and FDLs of musicians were less affected by noise than the FFRs and FDLs of non-musicians, two separate ANOVAs were calculated with factors group (between-subjects on 2 levels, musicians and non-musicians) and frequency (within-subjects on two levels, 500 and 1000 Hz). This was done by calculating the difference between quiet and noise FFR phase coherence, quiet and noise FFR amplitude, and quiet and noise FDLs.

Relationships between behavioral and physiological measures were assessed using Pearson-product moment correlations. Correlations were performed between each FDL measure and its corresponding condition for FFR phase coherence, between FDL measures and HINT scores, and between FFR phase coherence and HINT scores; Bonferroni corrections were applied to each set of correlations. To examine the effect of the number of years of musical training, Pearson-product moment correlations were also done between years of training and the behavioral and physiological measures.

Chapter IV

Results

Frequency Discrimination.

A three-factor ANOVA was conducted on FDLs. Factors were Group (between-subjects on 2 levels, musicians and non-musicians), frequency (within-subjects on two levels, 500 and 1000 Hz), and noise (within-subjects on two levels, quiet and noise). The main effects for group, F (1,26) = 11.59, p = .002, partial η^2 = .308, frequency F(1,26) = 7.077, p = .013, partial η^2 = .214, and noise F(1,26), = 15.30, p = .001, partial η^2 = .370, were significant. No interactions were significant ($p > .05$). Effect size was large for all main effects. Figure 3 summarizes the FDL data which shows that the FDLs were better in musicians, and that in both groups the FDLs were poorer in noise and poorer at 500 Hz.

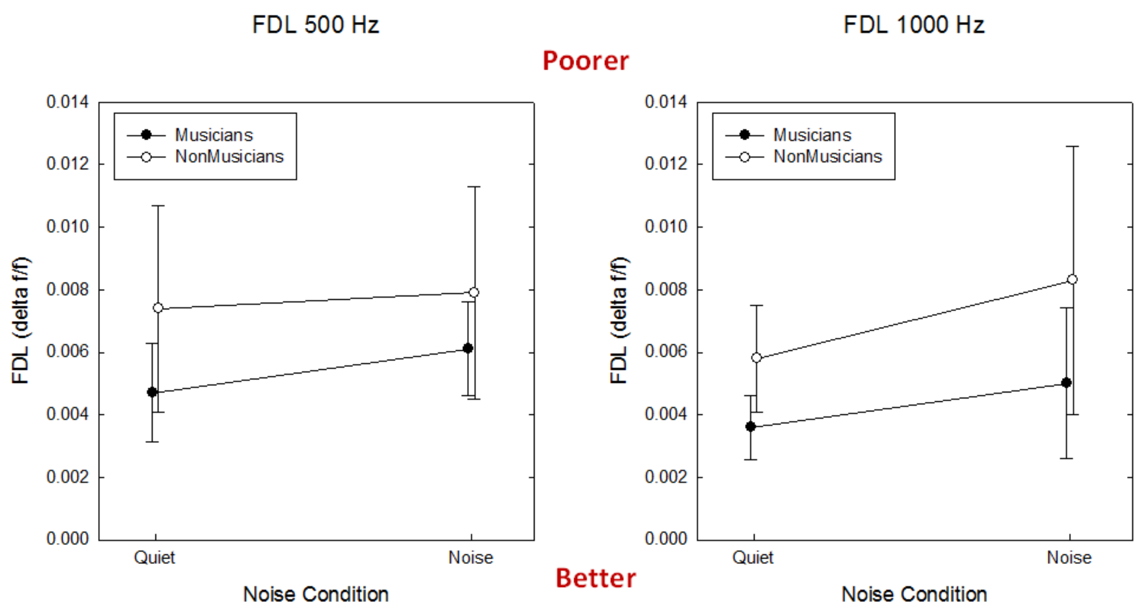

Figure 3. FDLs are shown by frequency, group, and noise conditions. Error bars represent one standard deviation.

Frequency-Following Response.

A three-factor ANOVA was conducted on FFR phase coherence. Factors were Group (between-subjects on 2 levels, musicians and non-musicians), frequency (within-subjects on two levels, 500 and 1000 Hz), and noise (within-subjects on two levels, quiet and noise). The main effect for frequency, F (1,26) = 16.28, p < .001, partial η^2 = .385, was significant, with a large effect size. The main effects for group, F(1,26) = 0.24, p = .625, partial η^2 = .009, and noise F(1,26) = 1.68, p = .207, partial η^2 = .061, were not significant (see Figure 3). Main effects for group and noise were small and medium, respectively. However, the group by noise interaction was significant, F (1,26) = 4.37, p =.047, partial η^2 = .144 (large effect size); for the non-musicians, phase coherence at 500 Hz was slightly improved in the noise condition (see figure 3).

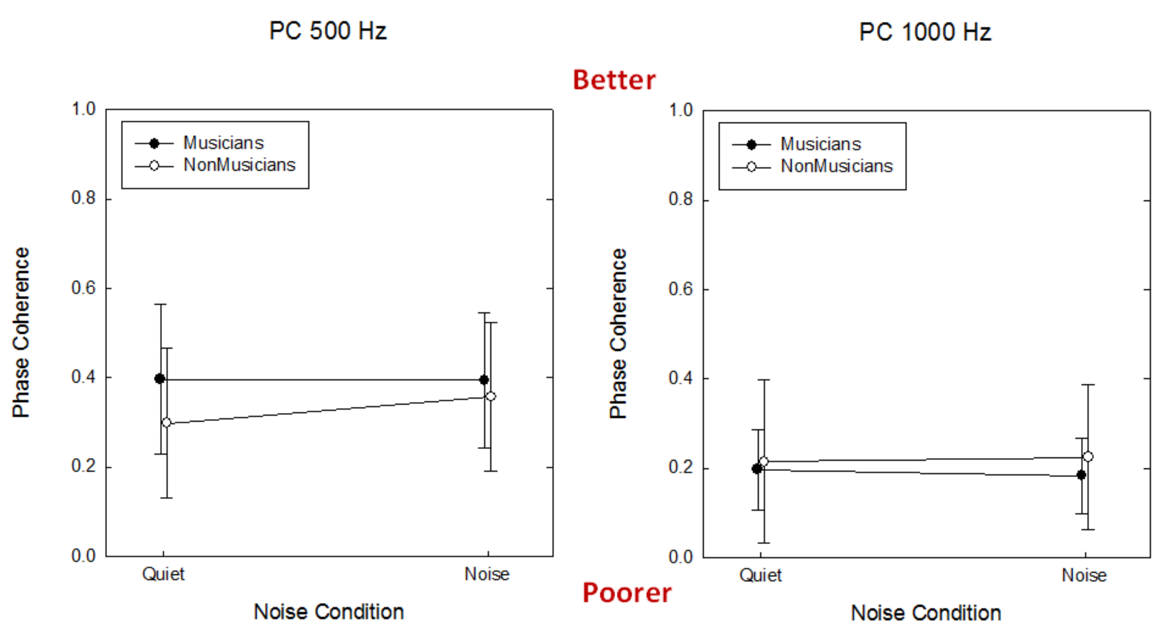

Figure 4. FFR phase coherence shown by frequency, group, and noise conditions. Error bars represent one standard deviation.

An additional three-factor ANOVA was conducted on FFR amplitude. Factors were Group (between-subjects on 2 levels, musicians and non-musicians), frequency (within-subjects on two levels, 500 and 1000 Hz), and noise (within-subjects on two levels, quiet and noise). The main effect for frequency, F (1,26) = 35.48, p < .001, partial η^2 = .577, was significant, indicating a large effect size. The main effects for group, F(1,26) = 0.11, p = .740, partial η^2 = .004, and noise F(1,26), = 0.19, p = .666, partial η^2 = .007, were not significant and indicated less than a small effect size. No interactions were significant (p > .05). These findings indicate that amplitudes for 1000 Hz were smaller than 500 Hz, but FFR amplitude was not significantly different across groups or noise conditions.

Speech-in-Noise.

In order to determine whether there was a significant difference in performance between musicians and non-musicians on the HINT, a one-way ANOVA was done with HINT SNRs as the dependent variable, and group as the factor (between-subjects on 2 levels, musicians and non-musicians). The signal-to-noise ratios (SNRs) from the HINT correspond to the 50% correct point. The results indicate no significant group differences, F(1,26) = 1.209, p = .282. The speech-in-noise understanding of musicians, as assessed by this measure, was not significantly better than non-musicians (see Figure 4).

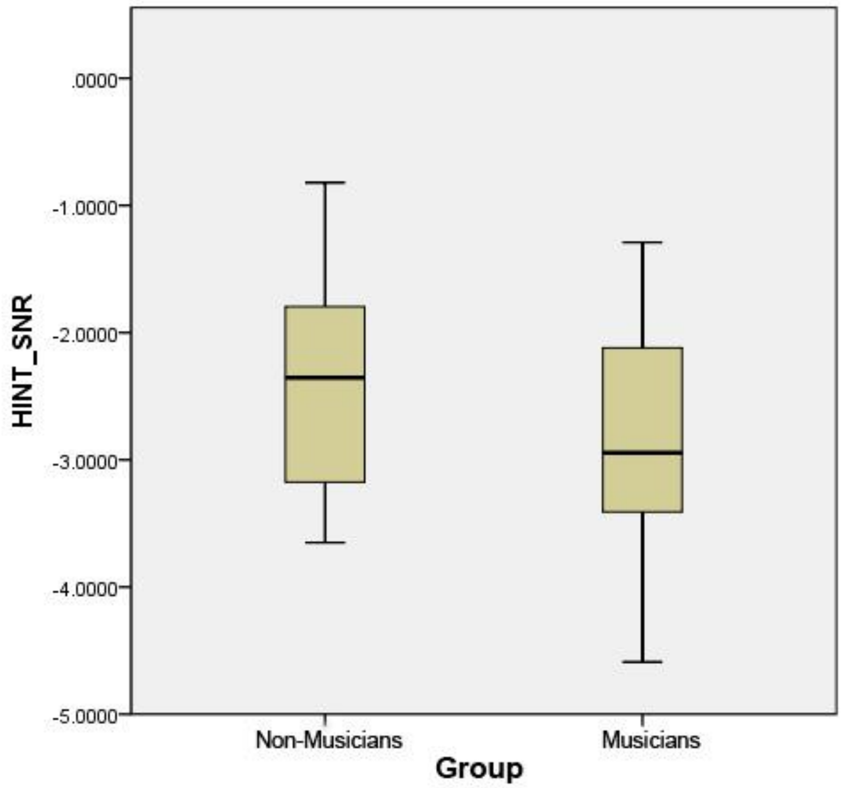

Figure 5. Box plot showing HINT Signal-to-noise ratios (SNRs) for musicians and non-musicians.

Relationships between measures.

Relationships between behavioral and physiological measures were assessed using Pearson-product moment correlations. For example, one of the behavioral frequency discrimination conditions would be paired with its corresponding physiological condition, such as FDLs obtained to 500 Hz in quiet and the 500 Hz in quiet FFR phase coherence. Bonferroni corrections were used for each set of correlations (e.g., 0.05/4 = .0125), bringing the p-value for a significant correlation to $p = 0.0125$.

FDL x FFR correlations.

The relationship between neural representation of TFS and the perception of TFS was examined using FDLs and the corresponding FFR condition. Correlations between the 500 Hz conditions in quiet (r = -.272, p = .161) and noise (r = -.009, p = .963) were not statistically significant. The correlation between the 1000 Hz conditions in quiet (r = .467, p = .012) was significant, but the correlation for the 1000 Hz conditions in noise(r = .403, p = .034) was not statistically significant. For the 1000 Hz in quiet conditions, lower (better) FDLs were significantly associated with lower (poorer) FFR phase coherence at 1000 Hz, which is opposite of the expected relationship. It had been hypothesized that lower FDLs would be related to higher phase coherence values. Scatter plots in Figure 5 illustrate these relationships.

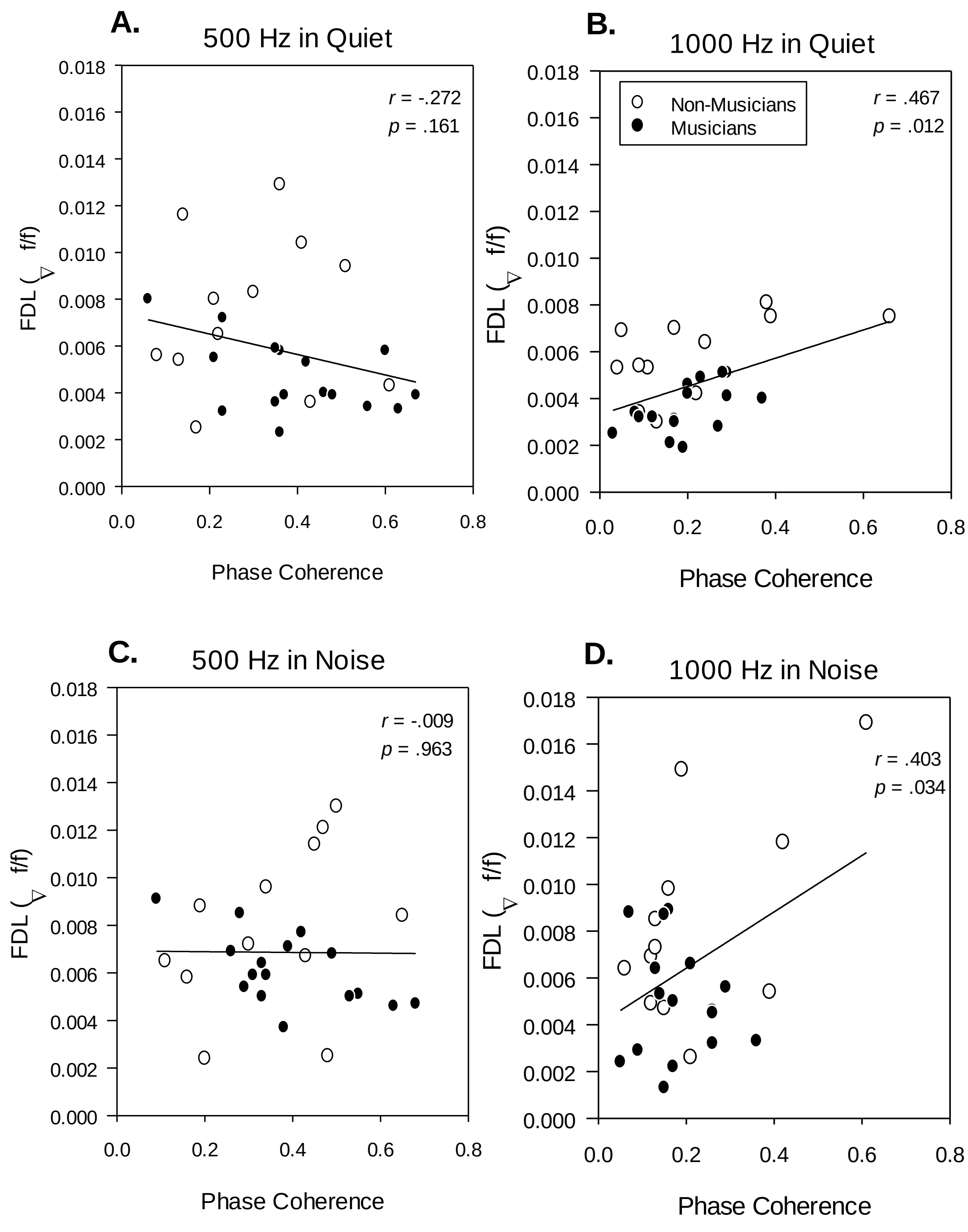

Figure 6. Scatter plots showing the relationship between behavioral FDLs and FFR phase coherence at 500 Hz in quiet and noise (A, C) and 1000 Hz in quiet and noise (B, D). Open symbols are non-musicians and filled symbols are musicians. Linear fit, correlation coefficient and p-value are shown for each panel and represent calculations based on all data shown, rather than one individual group.

FDL x HINT correlations

To assess the relationship between the perception of TFS and performance on a speech-in-noise measure, correlations between the FDL conditions and the HINT SNR were calculated. It was expected that lower (better) HINT SNRs would be associated with lower (better) FDLs. Correlations between HINT SNRs and 500 Hz FDLs in quiet (r = .415, p = .028) and in noise (r = .389, p = .041) were not statistically significant (Figure 6). Correlations between HINT SNRs and 1000 Hz FDLs in quiet (r = .338, p = .078) and in noise (r = .320, p = .097) were not statistically significant (Figure 6).

FFR x HINT correlations.

To assess the relationship between the neural representation of TFS and performance on a speech-in-noise measure, correlations between FFR phase coherence and HINT SNRs were calculated. It was expected that lower (better) HINT SNRs would be associated with higher (better) FFR phase coherence. Correlations between HINT SNRs and 500 Hz FFR phase coherence in quiet (r = -.350, p = .068) and in noise (r = -.438, p = .020) were not statistically significant (Figure 7). Correlations between HINT SNRs and 1000 Hz FFR phase coherence in quiet (r = .303, p = .117) and in noise (r = .266, p = .170) were not statistically significant (Figure 7).

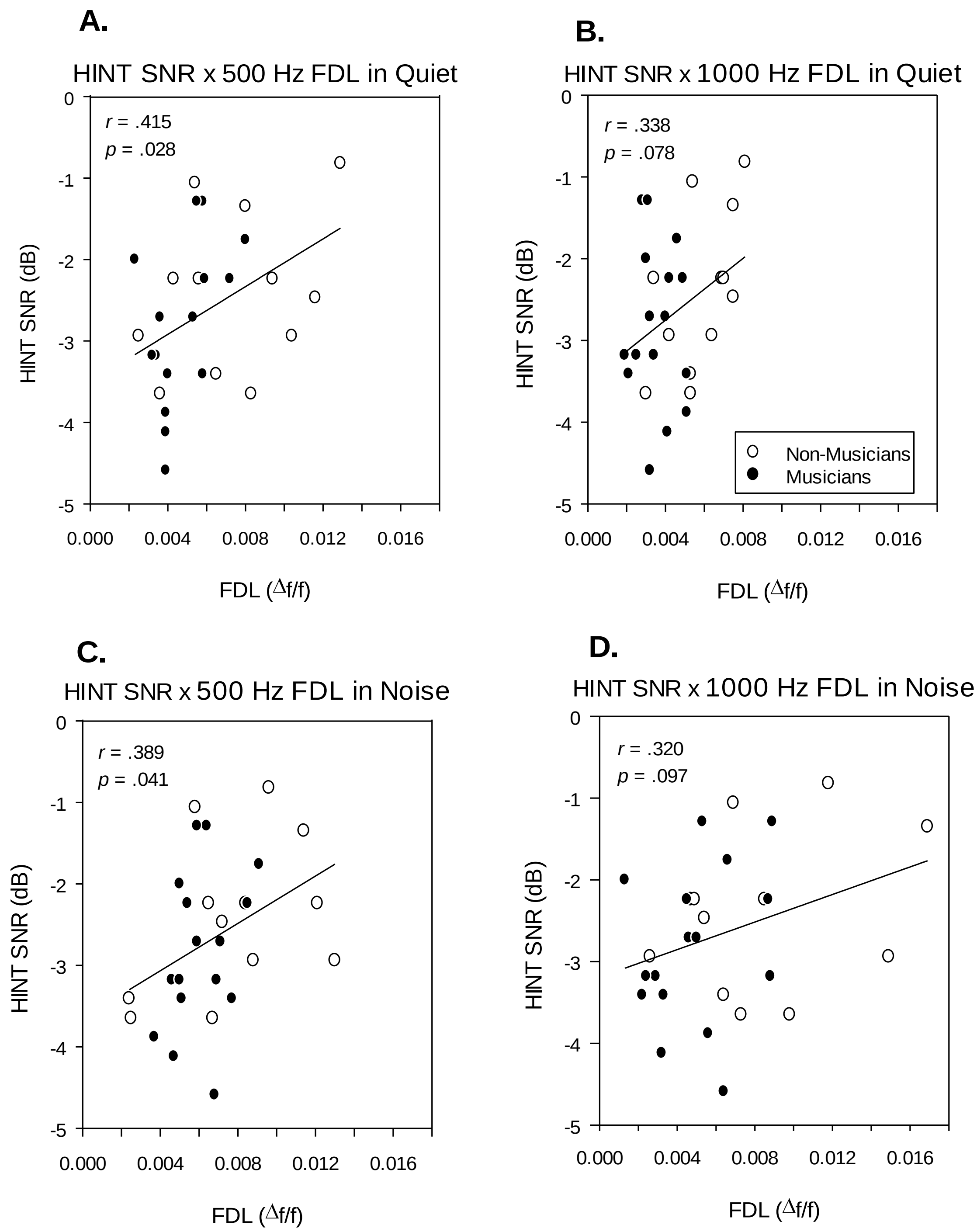

Figure 7. Scatter plots showing the relationship between HINT SNRs and behavioral FDLs at 500 Hz in quiet and noise (A, C) and 1000 Hz in quiet and in noise (B, D). Open symbols are non-musicians and filled symbols are musicians. Linear fit, correlation coefficient and p-value are shown for each panel and represent calculations based on all data shown, rather than one individual group.

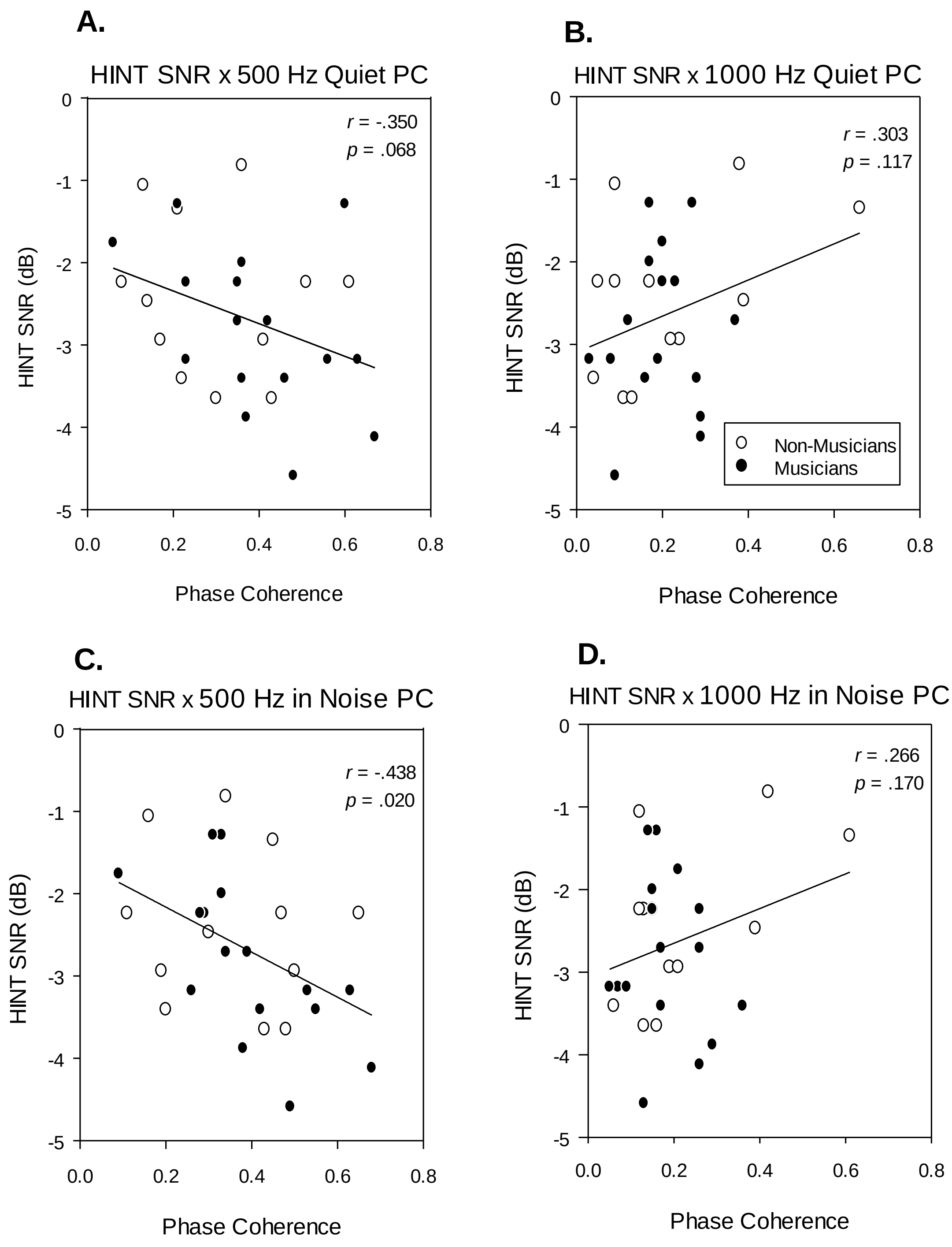

Figure 8. Scatter plots showing the relationship between HINT SNRs and FFR phase coherence at 500 Hz in quiet and noise (A, C) and 1000 Hz in quiet and noise (B, D). Open symbols are non-musicians and filled symbols are musicians. Linear fit, correlation coefficient and p-value are shown for each panel and represent calculations based on all data shown, rather than one individual group.

Years of musical training x measures.

Relationships between the number of years of musical training and the behavioral and physiological measures were assessed using Pearson-product moment correlations. Correlations were statistically significant for the FDLs conditions in quiet (see Table 1). For the significant correlations, a greater number of years was associated with lower (better) FDLs (see Figure 8). Correlations between years of musical training and FFR phase coherence were not statistically significant (see Table 2). Similarly, correlations between years of musical training and FFR amplitude were not statistically significant (Table 3). The number of years of musical training was also correlated with HINT SNRs and no statistical significance was found, $r = -.225$, $p = .251$. Bonferroni corrections were used for each set of correlations (e.g., 0.05/4 = .0125), bringing the p-value for a significant correlation to $p = 0.0125$.

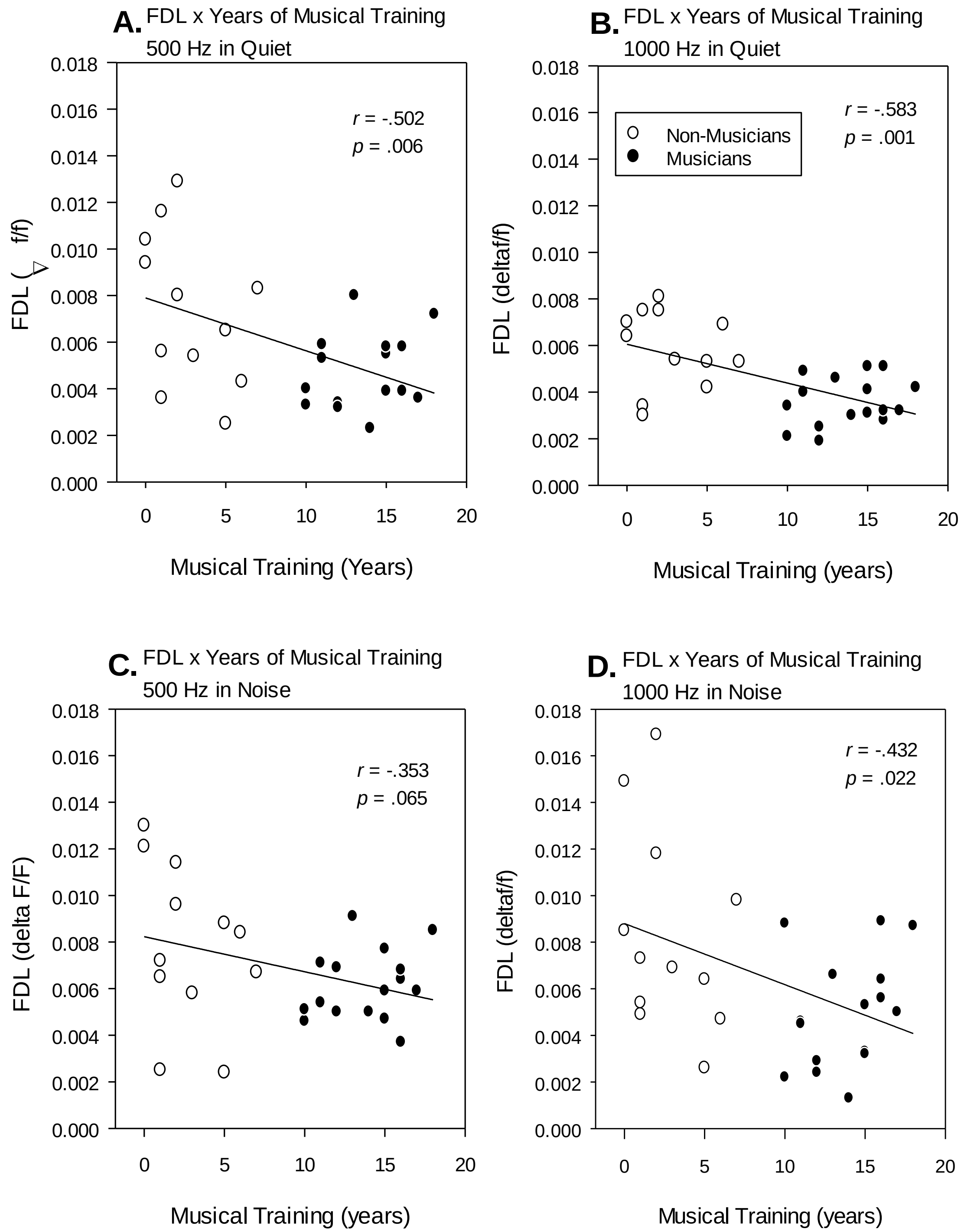

Figure 9. Scatter plots showing the relationship between years of musical experience and behavioral FDLs at 500 Hz in quiet and noise (A, C) and 1000 Hz in quiet and noise (B, D). Open symbols are non-musicians and filled symbols are musicians. Linear fit, correlation coefficient and p-value are shown for each panel.

Table 1. Pearson-product moment correlations between years of musical training and FDLs.

		FDL 500 Hz Quiet	FDL 1000 Hz Quiet	FDL 500 Hz Noise	FDL 1000 Hz Noise
Years Training	Pearson Correlation	-.502 **	-.583 **	-.353	-.432
	Sig. (2-tailed)	.006	.001	.065	.022
	N	28	28	28	28

** significant at the p = 0.0125 based on Bonferroni corrections

Table 2. Pearson-product moment correlations between years of musical training and FFR PC.

		PC 500 Hz Quiet	PC 1000 Hz Quiet	PC 500 Hz Noise	PC 1000 Hz Noise
Years Training	Pearson Correlation	.232	-.122	.061	-.175
	Sig. (2-tailed)	.234	.536	.759	.372
	N	28	28	28	28

Table 3. Pearson-product moment correlations between years of musical training and FFR Amp

		Amp 500 Hz Quiet	Amp 1000 Hz Quiet	Amp 500 Hz Noise	Amp 1000 Hz Noise
Years Training	Pearson Correlation	.115	-.027	.019	-.168
	Sig. (2-tailed)	.561	.891	.925	.392
	N	28	28	28	28

Resistance to noise.

To determine if behavioral FDLs or physiological FFRs of musicians were more resistant to effects of noise than non-musicians, differences were calculated between corresponding quiet and noise conditions (e.g., 500 Hz FDL in quiet – 500 Hz FDL in noise). Differences were calculated for FDLs and for FFR phase coherence, which were used as dependent variables. Two separate ANOVAs were calculated with factors of group (between-subjects on 2 levels, musicians and non-musicians) and frequency (within-subjects on two levels, 500 and 1000 Hz). The FDL difference ANOVA had non-significant main effects of group, F (1,26) = 0.015, *p* = .902, and frequency, *F* (1,26) = 1.81, *p* = .190; the interaction was also not significant (p > .05). The FFR difference ANOVA had a significant main effect of group, *F* (1,26) = 4.384, *p* = .046, but the main effect of frequency was not significant, *F* (1,26) = 2.694, *p* = .113, the interaction was not significant (p > .05). The significant main effect of group is likely driven by the non-musicians having a slight improvement in FFR phase coherence in noise at 500 Hz.

Chapter V

Discussion

The current study examined the effect of musical training on physiological and behavioral measures of temporal fine structure processing, as well as speech-in-noise understanding. The aims of the study were to examine the effects of musical training on the individual measures, in addition to the brain-behavior relationship between the measures. The hypotheses were 1) that musicians would demonstrate smaller (better) FDLs than non-musicians in quiet and in noise; 2) that musicians would have lower (better) HINT SNRs than non-musicians; 3) that musicians would demonstrate stronger FFR phase coherence and amplitude than non-musicians in quiet and in noise, and 4) that there would be significant relationships between behavioral and physiological measures. Lower (better) FDLs were expected to be associated with higher (better) FFR phase coherence, and vice versa. Lower (better) HINT SNRs were also expected to be associated with higher (better) FFR phase coherence. The results showed that musicians had significantly better FDLs than non-musicians, however no significant group effects were found for the FFRs or speech-in-noise understanding.

Musical experience on frequency discrimination.

The musicians displayed enhanced behavioral frequency discrimination compared to the non-musicians. It was hypothesized that musicians would have better (lower) FDLs than non-musicians in both quiet and in noise, and the results confirmed this. This is consistent with the literature on frequency discrimination in musicians (Kishon-Rabin et al., 2001; Micheyl et al., 2006; Nikjeh et al., 2008; Akin & Belgin, 2009). The results showed that performance on the 1000 Hz frequency discrimination task was significantly better than that of 500 Hz for both groups, and this was also an expected finding (Sęk & Moore, 1994; Micheyl et al., 2012). Furthermore it was expected that FDLs in quiet would be better than those in noise for both groups, and the results were consistent with this. Neither of the groups was significantly more affected by noise than the other. It was supposed that the non-musicians' FDLs would be more negatively affected by noise than the musicians. The results also showed that a greater number of years of musical experience was associated with better FDLs in quiet, but not in noise. Although it is not surprising that more experience would allow for better frequency discrimination in quiet, it would also be expected that it would improve frequency discrimination in noise, however the latter was not seen.

Musical experience and the FFR.

Significant findings for FFRs suggest that both groups had better phase coherence and amplitude measures for 500 Hz (quiet and noise) than 1000 Hz (quiet and noise). It is expected that phase-locking at lower frequencies is better than at high frequencies, so this is not a surprising finding (Clinard et al., 2010). There were no significant differences in FFR phase coherence by either group or between the quiet and noise conditions. This is contrary to what was hypothesized, because it was expected that musicians would have enhanced FFR measures over non-musicians and that performance in quiet would be better than in noise. FFRs of musicians have been reported to be more robust to the effects of noise than non-musicians (Parbery-Clark et al., 2009a).

It was also anticipated that a greater number of years of musical experience would heighten FFR measures (Wong et al., 2007; Parbery-Clark et al., 2011b), however this was not found. Previous FFR literature has reported that musicians have enhanced envelope/F_0 encoding (Musacchia et al., 2007; 2008; Wong et al., 2007). It is possible that musical experience may enhance phase-locked representations of fundamental frequency but experience-related changes do not affect phase-locked representations of TFS.

In order to compare this study's FFR findings with previous published literature, FFR phase coherence data of this present study was compared to data from Clinard et al. (2010) in Figure 10. Clinard et al. reported higher (better) phase coherence than the present study for both 500 Hz and 1000 Hz. However, it should be considered that the current study had 28 participants between the ages of 21-35, while Clinard et al. had 9 participants at 500 Hz and 7 participants at 1000 Hz. In the Clinard et al. study, age was used as a continuous variable, and so only a limited number of participants were from any given age range. Also, the toneburst duration in the present study was 300 ms, while it was 500 ms for Clinard et al. These differences should be taken into account when comparing data from the two studies.

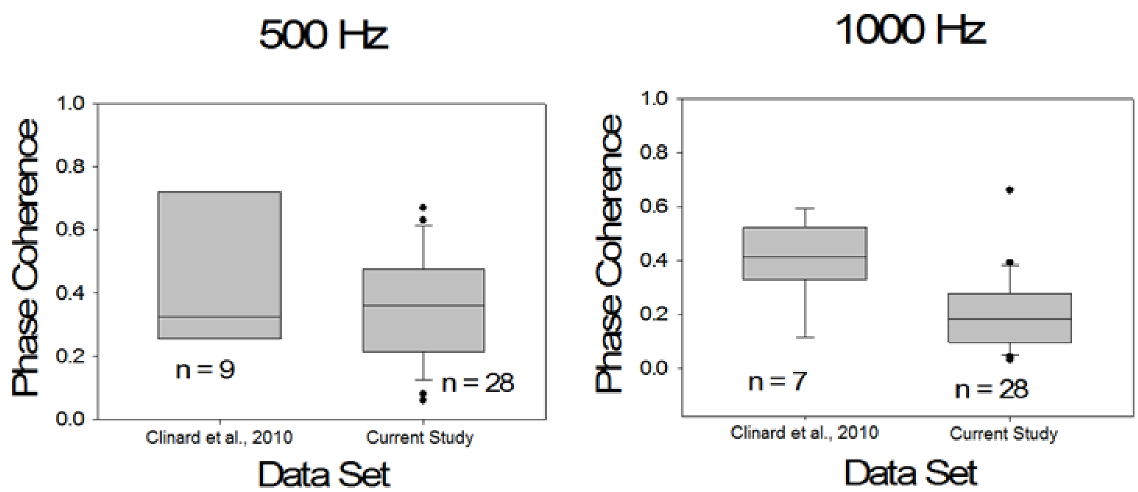

Figure 10. Box plots comparing FFR phase coherence data between Clinard et al., 2010 and the present study. The left panel shows data for 500 Hz in quiet, and the right panel shows data for 1000 Hz in quiet.

The group by noise interaction was significant which meant that for the non-musicians, phase coherence at 500 Hz was slightly improved in the noise condition. This is unexpected, and it may be related to test-retest variability of the FFR – perhaps the FFRs of non-musicians are less stable than those of musicians. It should however be noted that even though there was an improvement, the non-musicians still did not significantly perform better than the musicians in this condition.

Musical experience and Speech-in-Noise.

The results of the HINT SNRs indicate that the speech-in-noise understanding of musicians was not significantly better than non-musicians. This was an unexpected finding, as it had been hypothesized that musicians would have significantly better (more negative) SNRs than non-musicians. Multiple studies have reported that musicians can tolerate poorer SNRs than non-musicians (Parbery-Clark et al., 2009a; 2009b; 2011b; Soncini & Costa, 2006).

Another idea to be considered is that some studies (Parbery-Clark et al., 2011b) looking at the effects of musical experience on speech-in-noise testing used multitalker babble like the Quick Speech-in-Noise Test (QuickSIN), instead of steady noise with no dips in the envelope (e.g., HINT). It might be possible that the effects of musical training on speech-in-noise perception are more pronounced when a masker has dips in the noise (e.g., multitalker babble) that listeners can take advantage of (Parbery-Clark et al., 2009a; 2011b). Additionally, since language and music are different acoustic stimuli, it is suggested that using a speech test with low linguistic context may result in better comparisons between the various measures.

Surprisingly, there was also no significant effect of the number of years of musical training on HINT SNRs, which differs from data showing a negative correlation between QuickSIN SNRs and years of musical practice (Parbery-Clark et al., 2009a). However, a later study by Parbery-Clark and colleagues (Parbery-Clark et al., 2011b) showed no significant effect of years of musical practice on three different speech-in-noise tests (HINT, QuickSIN and WIN). It was suggested by this study that no significant correlations were found given the linearity of years of musical experience and the nonlinear nature of perceptual measures.

Brain-Behavior Relationships/Relationships between measures.

It was anticipated that the neural representation of TFS and the behavioral measures of TFS would be correlated in a way that showed better FFR responses to be associated with better FDLs. However, there was a lack of a predictive relationship between FDLs and FFR phase coherence except for 1000 Hz in quiet. For this one condition, lower (better) FDLs were significantly associated with lower (poorer) FFR phase coherence at 1000 Hz, which is the opposite trend from what was hypothesized.

In looking at the FDL x FFR figures it appears that the musicians follow the expected trend for the 500 Hz conditions (quiet and noise), where negative correlation coefficients mean that poorer FDLs are associated with poorer phase coherence (Figure 5, panels A and C). Again, these relationships were not statistically significant however.

Clinard and colleagues also did not find a predictive relationship between FFR and FDL measures (Clinard et al., 2010). It was suggested that the absence of an overall significant relationship between the two could be associated with the cues the subjects used during the behavioral task, i.e. in addition to or in place of temporally phase-locked representations of frequency, the participants may have also used place-based cues. Furthermore, the behavioral frequency discrimination employs several neural pathways and greater neural activity, whereas the FFR reflects one representation of frequency. The stimulus context of each of the measures is yet another factor that may be involved in the poor brain-behavior relationship. For the frequency discrimination task, the stimuli were presented as pairs of tones, however repeated presentations of a single tone were used to record the FFR, and the two require different sensory processing.

Additionally, it was expected that better speech-in-noise understanding would be related to better behavioral frequency discrimination as well as the neural representation of TFS. It was hypothesized that lower (better) HINT SNRs would be correlated with lower (better) FDLs, and lower (better) SNRs with higher (better) FFR phase coherence. This was not the case for any of the FDL conditions or FFR phase coherence. In contrast, few papers have reported correlations between speech-in-noise measures and F_0 amplitude (Parbery-Clark et al., 2011b; Strait, Parbery-Clark, Hittner, & Kraus, 2012). The lack of significant correlations with the speech-in-noise task may be, in part, attributed to the difference between stimulus contexts. For example, the HINT is a speech/language-based assessment, whereas the frequency discrimination and physiologic measures used tonal stimuli, and as such different types of auditory processing were required for these measures.

Methodological Issues.

Some of the methodological issues to consider would be subject selection criteria, and choice of background noise. For this study, the criteria for musicians were at least 10 years of consistent musical training, either formally or self-taught, professional or amateur. There are some studies which have suggested using classically trained or professional musicians (Spiegel & Watson, 1984; Kishon-Rabin, 2001). It is possible that more, or larger, group differences may have been found if the musicians were required to be professional or more highly-trained, with a greater number of years of experience. Nevertheless, in this study the rationale for the musicians' criteria was chosen because it is more realistic for the overall population of musicians and because of its implications for the general population, since each individual cannot become a professional musician in order to get the benefits of enhanced auditory processing. Even so, the task of trying to quantify or standardize the level of one's musicianship is not straightforward, and thus in this study, as with many studies looking at musicians, it is difficult to say whether the musicians were good musicians or not. Group differences on the FFR measures may have also be more apparent at less favorable SNRs. Instead of the

20 dB SNRs used in this study, more challenging SNRs (e.g., 5 dB) may be more effective at revealing group differences.

In this study, the background noise was steady-state, but the research shows a stronger dependence on TFS when the noise is modulated (Parbery-Clark et al., 2009b; 2011b). Multi-talker babble versus steady, speech-shaped noise may change the relationship and ability to listen in the dips of the masker, and allow thus the listener to take advantage of TFS. A speech-in-noise material that has less linguistic context than the HINT's sentences and a masker with amplitude variations (i.e., multi-talker babble) may be more effective at evaluating the relationship between speech-in-noise understanding and behavioral and physiological representations of TFS.

Clinical Applications.

One of the most common complaints of persons with hearing loss is the difficulty that is faced when listening to speech in the presence of background noise. Much research across the field is devoted to finding ways in which this can be improved. Musical training has been one of the recommendations for improved speech-in-noise understanding, and many studies have reported this (Parbery-Clark et al., 2009a; 2009b; 2011b; Soncini & Costa, 2006), along with other auditory benefits. Because of the evidence pointing to overall improvement of auditory listening tasks of musicians, it should certainly still be recommended that musical training is of benefit to the individual and would do no harm, regardless of the findings of this study. Parbery-Clark, Anderson, Hittner, and Kraus (2012) compared musicians and non-musicians in the 18-32 years range and the 46-65 years age range and found that the older musicians had youth-like FFR latencies, FFR F_0 amplitude, and speech-in-noise measures. In addition, Parbery-

Clark et al. (2011b) reported that older musicians had youth-like cognitive measures – working memory, reaction time, speech-in-noise understanding – compared to age-matched non-musicians. These data indicate that musical training or experience has the effect of slowing aging effects on specific functions of the CANS, as well as some cognitive processes. However, data from the present study indicate that CANS effects of musical training are not global in nature. Previous literature supports the idea that fundamental frequency encoding is enhanced by musical experience, but the present study shows no enhancement of TFS encoding with musical experience.

It is important to mention however that there is a question of whether musical training affects brain plasticity or whether there are certain people who are born with brain characteristics that predispose them towards being musicians (Strait et al., 2012). It is likely that both are true. We can see the effect of musical training on plasticity by the improvement of auditory tasks with years of musical experience. Also, it is often observed that certain families tend to be "musical families", whereas others are not.

Future Directions.

Many questions remain about the effects of musical training and how it can be utilized clinically. The interactions between hearing loss and musical training, and how this relates to behavioral and physiological measures of envelope, TFS, and speech-in-noise understanding should be examined. Also, the question of whether musical training later in life can produce auditory benefits that can be seen by improved speech in noise understanding and/or frequency discrimination and neural TFS is an area worth studying. Furthermore, it would be interesting to look at whether there are any auditory benefits of being an avid music listener, as opposed to actively being engaged in music making.

Since frequency discrimination inherently needs two or more stimuli for discrimination, it would be interesting to see whether absolute pitch ("perfect pitch") can be used as a behavioral measure of TFS, and whether a brain-behavior relationship can be found between the two.

Conclusions.

(1) Musicians had better FDLs than non-musicians

(2) Musical experience did not affect phase-locked representations of TFS.

(3) Musicians did not have better signal-to-noise ratios on the HINT, indicating no significant speech-in-noise advantage related to musical training.

(4) FFR phase coherence was not significantly related to FDLs or HINT SNRs, with the exception of a significant correlation between FFR phase coherence and FDLs for 1000 Hz in quiet.

(5) A greater number of years of musical experience was related to lower (better) FDL thresholds across all stimuli except 500 Hz in noise. The years of training did not relate to performance on FFR phase coherence, amplitude, or HINT scores.

References

Abel, S. M., Krever, E. M. & Alberti, P. W. (1990). Auditory detection, discrimination and speech processing in ageing, noise-sensitive and hearing-impaired listeners. *Scandinavian Audiology, 19,* 43-54.

Akin, O. & Belgin, E. (2009). Hearing characteristics and frequency discrimination ability in musicians and nonmusicians. *The Journal of International Advanced Otology, 5*(2), 195-202.

Apoux, F. & Healy, E. W. (2011). Relative contribution of target and masker temporal fine structure to the unmasking of consonants in noise. *Journal of the Acoustical Society of America, 130*(6), 4044-4052.

Bidelman, G. M., Gandour, J. T. & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition, 77,* 1-10.

Buss, E., Hall, J. W. & Grose, J. H. (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear & Hearing, 25*(3), 242-250.

Campbell, R. A. & Small, A. M. (1963). Effect of practice and feedback on frequency discrimination. *Journal of the Acoustical Society of America, 35*(10), 1511-1514.

Carcagno, S. & Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology, 12*(1), 89-100.

Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hearing Research, 264*, 48-55.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Psychology Press.

Demany, L. (1985). Perceptual learning in frequency discrimination. *Journal of the Acoustical Society of America*, *78*(3), 1118-1120.

Dobie, R. A. & Wilson, M. J. (1989). Analysis of auditory evoked potentials by magnitude-squared coherence. *Ear & Hearing*, *10*(1), 2-13.

Drennan, W. R., Won, J. H., Dasika, V. K. & Rubinstein, J. T. (2007). Effects of temporal fine structure on the lateralization of speech and on speech understanding in noise. *Journal of the Association for Research in Otolaryngology*, *8,* 373-383.

Draganova, R., Wollbrink, A., Schulz, M., Okamoto, H. & Pantev, C. (2009). Modulation of auditory evoked responses to spectral and temporal changes by behavioral discrimination training. *BioMed Central Neuroscience*, *10*(143).

Dye, R. H., Jr. & Hafter, E. R. (1980). Just-noticeable differences of frequency of masked tones. *Journal of the Acoustical Society of America*, *67*(5), 1746-1753.

Fisher, N. I. (1993). *Statistical analysis of circular data.* New York, NY: Cambridge University Press.

Geiser, E., Sandmann, P., Jäncke, L. & Meyer, M. (2010). Refinement of metre perception training increases hierarchical metre processing. *European Journal of Neuroscience, 32*, 1979-1985.

Geringer, J. M. (1995). Continuous loudness judgments of dynamics in recorded music excerpts. *Journal of Research in Music Education, 43*(1) 22-35.

Grose, J. H. & Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear & Hearing, 31*(6), 755-760.

Herholz, S. C., Lappe, C. & Pantev, C. (2009). Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BioMed Central Neuroscience, 10*(42).

Hienz, R. D., Sachs, M. B. & Aleszczyk, C. M. (1992). Frequency discrimination in noise: Comparison of cat performances with auditory-nerve models. *Journal of the Acoustical Society of America, 93*(1), 462-469.

Hopkins, K. & Moore, B. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *Journal of the Acoustical Society of America, 130*(1), 334-349.

Humes, L. E. (1996). Speech understanding in the elderly. *Journal of the American Academy of Audiology, 7,* 161–167.

Irvine, D.R.F., Martin, R. L., Klimkeit, E. & Smith, R. (2000). Specificity of perceptual learning in a frequency discrimination task. *Journal of the Acoustical Society of America, 108*(6), 2964-2968.

John, M. S., Lins, O. G., Boucher, B. L. & Picton, T. W. (1998). Multiple auditory steadystate responses (MASTER): stimulus and recording parameters. *Audiology 37*(2), 59-82.

Kishon-Rabin, L., Amir, O., Vexler Y. & Zaltz Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic & Clinical Physiology & Pharmacology*, *12*(2), 125-143.

Koelsch, S., Schröger, E. & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *NeuroReport, 10,* 1309-1313.

Kraus, N., Skoe, E., Parbery-Clark, A. & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre, and timing: Implications for language and music. *The Neurosciences and Music III—Disorders and Plasticity: Annals of the New York Academy of Sciences, 1169*, 543–557.

Krishnan, A. (2007). Frequency-following response. In Burkhard, R. F., Don, M. & Eggermont, J. J. (Eds.), *Auditory Evoked Potentials: Basic principles and clinical applications* (313-333). Baltimore, MD: Lippincott Williams & Wilkins

Krishnan, A., Gandour, J. T. & Bidelman, G. M. (2010). Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. *Brain Research 1313,* 124-133.

Krishnan, A., Gandour, J. T. & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics, 23,* 81-95.

Krishnan, A., Gandour, J. T., Bidelman, G. M. & Swaminathan, J. (2009). Experience-dependent neural representation of dynamic pitch in the brainstem. *NeuroReport, 20,* 408-413.

Krishnan, A., Xu, Y., Gandour, J. T. & Cariani, P. A. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research 25*, 161-168.

Lee, K. M., Skoe, E., Kraus, N. & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *The Journal of Neuroscience, 29*(18), 5832–5840.

Levi, E. C., Folsom, R. C. & Dobie, R. A. (1995). Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. *Hearing Research, 89,* 21-27.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49,* 467-477.

Micheyl, C., Delhommeau, K., Perrot, X. & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research, 219,* 36-47.

Micheyl, C., Moore, C. B. J. & Carlyon, R. P. (1998). The role of excitation-pattern cues and temporal cues in the frequency and modulation-rate discrimination of amplitude-modulated tones. *Journal of the Acoustical Society of America, 104*(2), 1039-1050.

Micheyl, C., Xiao, L. & Oxenham, A. J. (2012). Characterizing the dependence of pure-tone frequency difference limens on frequency, duration, and level. *Hearing Research, 292,* 1-13.

Miyazono, H., Glasberg, B. R. & Moore, B. C. J. (2010). Perceptual learning of fundamental frequency discrimination: Effects of fundamental frequency, harmonic number, and component phase. *Journal of the Acoustical Society of America, 128*(6), 3649–3657.

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.

Musacchia, G., Strait, D. & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research, 241,* 34-42.

Neher, T., Lunner, T., Hopkins, K. & Moore, B. (2012). Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners. *Journal of the Acoustical Society of America, 131*(4), 2561-2564.

Nikjeh, D. A., Lister, J. J. & Frisch, S. A. (2008). Hearing of note: An electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *Psychophysiology, 45*, 994–1007.

Olsho, L. W., Koch, E. G. & Carter, E.A. (1988). Nonsensory factors in infant frequency discrimination. *Infant Behavior and Development, 11,* 205-222.

Ott, C.G., Langer, N., Oechslin, M., Meyer, M. & Jäncke, L. (2011). Processing of voiced and unvoiced stimuli in musicians. *Frontiers in Psychology, 2*(195), 1-10.

Oxenham, A. J., Fligor, B. J., Mason, C. R. & Kidd, J. (2003). Informational masking and musical training. *Journal of the Acoustical Society of America, 114*(3), 1543-1549.

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto1, J., Gjedde, A. & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS ONE, 5*(6), 1-12.

Parbery-Clark, A., Anderson, S., Hittner, E. & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging, 33,* 1483.e1-1483.e4.

Parbery-Clark, A., Marmel, F., Bair, J. & Kraus, N. (2011). What subcortical-cortical relationships tell us about processing speech in noise. *European Journal of Neuroscience, 33*, 549-557.

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.

Parbery-Clark, A., Skoe, E., Lam, C. & Kraus, N. (2009). Musician enhancement in speech-in-noise. *Ear & Hearing, 30*(6), 653-661.

Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E. & Kraus, N. (2011). Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *PLoS ONE 6*(5), 1-8.

Plack, C. J., Turgeon, M., Lancaster, S., Carlyon, R. P. & Gockel, H. E. (2011). Frequency discrimination duration effects for Huggins pitch and narrowband noise (L). *Journal of the Acoustical Society of America, 129*(1), 1-4.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions: Biological Sciences, 336*(1278), 367-373.

Seither-Preisler, A., Johnson, L., Krumbholz, K., Nobbe, A., Patterson, R., Seither, S. & Lütkenhöner, B. (2007). Tone sequences with conflicting fundamental pitch and timbre changes are heard differently by musicians and nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 743-751.

Sęk, A. & Moore, B. C. J. (1994). Frequency discrimination as a function of frequency, measured in several ways. *Journal of the Acoustical Society of America, 97*(4), 2479-2486.

Sęk, A. & Moore, B. C. J. (2012). Implementation of two tests for measuring sensitivity to temporal fine structure. *International Journal of Audiology, 51,* 58-63.

Smith, J. C., Marsh, J. T., & Brown, W. S. (1975). Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology, 39,* 465-472.

Soncini, F. & Costa, M. J. (2006). The effect of musical practice on speech recognition in quiet and noisy situations. *Pro Fono, 18*(2), 161-170.

Song, J. H., Skoe, E., Wong, P. C. M. & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience, 20*(10), 1892-1902.

Spiegel, M. F. & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. *Journal of the Acoustical Society of America, 76*(6), 1690-1695.

Strait, D. L., Parbery-Clark, A., Hittner, E. & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain & Language, 123,* 191-201.

Strelcyk, O. & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *Journal of the Acoustical Society of America, 125*(5), 3328-3345.

Swaminathan, J., Krishnan, A., Gandour, J. T. & Xu, Y. (2008). Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Transactions on Biomedical Engineering, 55*(1), 281-287.

Tervaniemi, M. Just, V., Koelsch, S., Widmann, A. & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental Brain Research, 161*, 1-10.

Tervaniemi, M., Kruck, S., De Baene, W., Schröger, E., Alter, K. & Friederici, A. D. (2009). Top-down modulation of auditory processing: effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience, 30,* 1636-1642.

Tierney, A. T., Bergeson, T. R., & Pisoni, D. B. (2008). Effects of early musical experience on auditory sequence memory. *Empirical Musicology Review, 3*(4), 178-186.

Trainor, L. J., Desjardins, R. J. & Rockel, C. (2011). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Australian Journal of Psychology, 51*(3), 147-153.

Tremblay, C., Picard, M., Barbarosie, T. & Banville, R. (1991). Clinical study of speech understanding in noise. *Audiology, 30*(4), 212-240.

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C. & Roepstorff, A. (2009). Predictive coding of music-Brain responses to rhythmic incongruity. *Cortex, 45,* 80-92.

Wang, S., Staffaroni, L., Reid, E. Jr., Steinschneider, M. & Sussman, E. (2009). Effects of musical training on sound pattern processing in high-school students. *International Journal of Pediatric Otorhinolaryngology, 73,* 751-755.

Wang, S., Xu, L. & Mannell, R. (2011). Relative contributions of temporal envelope and fine structure cues to lexical tone recognition in hearing-impaired listeners. *Journal of the Association for Research in Otolaryngology, 12,* 783-794.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T. & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.

Wright B. A. & Ortiz, J. A. (2000). Unexpected benefits of training frequency discrimination on the learning and generalization of interval discrimination. *Journal of the Acoustical Society America, 107*(5), 2821.

Wright, B. A. (2001). Why and how we study human learning on basic auditory tasks. *Audiology & Neuro-Otology, 6*, 207-210.

Zendel, B. R. & Alain, C. (2011). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging,* 1-8.