

MUSICIANSHIP AND NEURAL SYNCHRONIZATION
AT MULTIPLE TIMESCALES

by

Gabriel A. Nespoli

B.Sc, Psychology, McGill University, 2007

A thesis presented to

Ryerson University

submitted in partial fulfillment of the

requirements for the degree of

Master of Arts

in the Program of

Psychology

Toronto, Ontario, Canada, 2014

© Gabriel A. Nespoli, 2014

AUTHOR'S DECLARATION FOR
ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

Abstract

MUSICIANSHIP AND NEURAL SYNCHRONIZATION AT MULTIPLE TIMESCALES

Master of Arts, 2014

Gabriel A. Nespoli

Psychology, Ryerson University

Auditory events can be considered to have spectral energy at short and long timescales, corresponding to the musical phenomena of pitch and pulse. Neural synchronization—when neurons synchronize their firing with external oscillatory stimuli—can be measured using spectral EEG at both subcortical and cortical levels. It has been shown that subcortical synchronization to tones is more robust in musicians than nonmusicians, suggesting a type of experience-dependent plasticity; a similar test for long timescales has not been investigated. In the current study, EEG was measured from musicians and nonmusicians while they listened to an isochronous sequence of tones. Neural synchronization at short timescales was found to be stronger in musicians. Additionally, the extent of synchronization correlated with the current level of musical engagement. These findings indicate that the experience-dependent plasticity observed in musicians manifests itself at multiple cortical levels corresponding to oscillations at different timescales present in music.

Acknowledgements

I would like to express my wholehearted gratitude to all of my friends and family who have supported me both academically and musically. I am especially grateful to my supervisor, Frank Russo, who provided me the opportunity to work in the lab and has helped to make the transition to graduate school an enjoyable and invigorating one. His guidance and support continues to be invaluable. Many thanks to Paolo Ammirante for countless discussions and insightful comments throughout the project. To all the members of the SMART Lab, past and present, who have helped create a context of rousing discussion and lasting friendship, I am thankful to be a part of such an extraordinary group of individuals. And to Erin Nespoli, without whose constant encouragement and belief I could not have come so far.

Table of Contents

| | |
|--|----|
| Musicianship and Neural Synchronization at Multiple Timescales | 1 |
| Multiple Timescales: Acoustics..... | 2 |
| Multiple Timescales: Sensorimotor Synchronization..... | 5 |
| Multiple Timescales: Neural Correlates | 7 |
| Subcortical Synchronization at Short Timescales..... | 7 |
| Subcortical Synchronization and Musicianship | 8 |
| Cortical Synchronization at Long Timescales..... | 9 |
| Cortical Synchronization and Musicianship..... | 11 |
| Resonance..... | 11 |
| Musicianship and Multiple Timescales..... | 13 |
| Methods..... | 15 |
| Participants..... | 15 |
| Stimuli and Procedure | 16 |
| Analysis and Results..... | 21 |
| Discussion..... | 24 |
| Musicianship and Neural Synchronization..... | 24 |
| Synchronization Across Multiple Timescales | 28 |
| Complex Periodicities..... | 29 |
| Higher-Level Synchronization Processes | 31 |
| Limitations..... | 32 |
| Conclusion..... | 33 |
| Tables | 34 |
| Figures | 36 |
| References | 45 |

List of Tables

Table 1. Independent samples t-tests on spectral peaks (subcortical and cortical) between musicians and nonmusicians34

Table 2. Correlations between neural measures of synchronization (subcortical and cortical) and behavioural measures of musicianship35

List of Figures

| | |
|---|----|
| <i>Figure 1a.</i> Single-cell recordings from various auditory structures in the cat to a tone stimulus | 36 |
| <i>Figure 1b.</i> Recordings from the vertex of the scalp to a tone stimulus from a cat and a human | 37 |
| <i>Figure 2a.</i> Waveform of the tone stimulus | 38 |
| <i>Figure 2b.</i> Amplitude spectrum of the stimulus | 38 |
| <i>Figure 3a.</i> Frequency following responses for musicians and nonmusicians | 39 |
| <i>Figure 3b.</i> FFR amplitude spectra for musicians and nonmusicians | 39 |
| <i>Figure 4.</i> SSEP amplitude spectra for musicians and nonmusicians | 40 |
| <i>Figure 5a.</i> Subcortical synchronization at f_0 of the tone stimulus. | 41 |
| <i>Figure 5b.</i> Cortical synchronization at the frequency of the pulse | 41 |
| <i>Figure 6a.</i> Subcortical synchronization to harmonics of the tone | 42 |
| <i>Figure 6b.</i> Cortical synchronization to the pulse and related frequencies | 42 |
| <i>Figure 7a.</i> Spectral FFR peaks correlate with hours playing music | 43 |
| <i>Figure 7b.</i> Spectral SSEP peaks are marginally correlated with hours spent listening to music | 43 |
| <i>Figure 8.</i> Neural synchronization at short and long timescales predicts musical engagement | 44 |

Musicianship and Neural Synchronization at Multiple Timescales

Audition is a temporal phenomenon because it involves interpreting the way in which the auditory display unfolds over time. Sounds create an oscillatory pressure wave through the air that resonates the eardrum, and it is the timing of the eardrum's oscillations that give rise to certain aspects of the perception of a sound. For example, the frequency of a pure tone is what gives rise to its unique pitch, and the time interval between note onsets is what gives rise to its unique tempo. Most sounds contain many periodicities that combine together, where phase and energy relationships between frequencies further define what it sounds like. A complex tone typically contains frequencies that are related to the slowest periodicity in the tone—the fundamental frequency—by simple integer ratios, and the relative loudness of all the frequencies give rise to its timbre. In a sequence of musical tones, such as a melody, the pattern and distribution of fundamental frequencies generates a tonal context that determines the stability and salience of tones that follow (Krumhansl & Kessler, 1982). Differences in tonal context can also convey unique affective experiences. For example, major and minor keys are tonal contexts that differ with regard to only one or two fundamental frequencies, yet major modes tend to convey happiness while minor modes tend to convey sadness (Dalla Bella, Peretz, Rousseau, & Gosselin, 2001; Hevner, 1935; Webster & Weir, 2005).

A second level of periodicity in most music occurs at a timescale that is slower than the periodicity found in tones. In particular, the timing of the onsets of successive tones tends to be quasi-periodic and can define how other aspects of a tone sequence are perceived. For example, an isochronous tone sequence contains tones that are sounded at a constant rate, whereas a complex rhythm contains multiple rates that combine together in much the same way as tones. Like tones, different rates have been associated with unique affective

experiences, as fast rates tend to convey happiness while slow rates tend to convey sadness (Dalla Bella et al., 2001; Webster & Weir, 2005).

These timing patterns of tones and rhythms—or periodicities at short and long timescales, respectively—are defining characteristics of music and sound. The intricate relations between periodicities give rise to perceptions of pitch and rhythm, and can influence emotion. Interacting with these periodicities is something that humans are readily able to do as is evidenced by group singing, conversation, and dance. How we come to perceive these periodicities in music has been the target of extensive research, but many questions remain. The current thesis takes aim at the timing elements of music, focusing on how periodicities at multiple timescales come to be represented via neural oscillations. In the sections that follow, multiple timescales are discussed from the perspective of acoustics, human behaviour, and neural oscillation.

Multiple Timescales: Acoustics

An auditory stimulus can be characterized as having levels of temporal structure at different timescales. Pitch, for example, is the phenomenological result of temporal regularities or periodicities at short timescales, involving periods that are typically shorter than 50 ms (i.e., 20 Hz or more). The onset of each period is not perceivable, however, so the entire oscillation is perceived as a single event. Rhythm or pulse refers to the organization of many events, occurring at long timescales with periodicities longer than 250 ms (i.e., 4 Hz or less). Since the onset of each oscillation is noticeable to the listener they are perceived as distinct events, and this sequence of events as a whole can be perceived as a rhythm.

In complex sounds, many periodicities combine together to create rich auditory displays, but there is often a single periodicity that dominates perception. For tones (short timescale), this is the fundamental frequency (f_0)

and corresponds to the slowest oscillation present in the tone. All the frequencies present in a single tone tend to be related by simple integer ratios to f_0 such as 1:2 or 2:1, called harmonics. For example, a complex tone with a fundamental frequency of 100 Hz would have 200 Hz (1:2) as its first harmonic, 300 Hz (1:3) as its second harmonic, and so on. It is with f_0 that we associate pitch; a 262 Hz tone would correspond to C₄ or “middle C” (approximately) on a piano keyboard. A tone with a faster f_0 such as D₄ (294 Hz) would sound “higher” in pitch. The relative amplitude of each harmonic within a tone give rise to the perception of timbre. For musical rhythms (long timescale), the periodicity that dominates perception is called the pulse or beat. Although the pulse is the periodicity at which listeners will feel most comfortable tapping, subdivisions and groupings of the pulse may be implied by the music and/or imposed by the listener. These subdivisions and groupings, like harmonics in their relation to f_0 , also tend to form simple integer ratios with the pulse frequency. Typical subdivisions of the pulse are duple or triple, where a each pulse may be divided into two or three equal parts (forming ratios of 1:2 or 1:3). Typical groupings of the pulse are also duple or triple, where successive pulses are perceived in groups of two or three (forming ratios of 2:1 or 3:1).

There are individual differences in the ability to discriminate auditory events based on their periodicity, with smaller discrimination thresholds tending to correlate with experience. Musicians, for example, interact with pitch discriminations every time they play or listen to music, since they are interested in discerning the subtle relationships between them. Micheyl et al. (2006) tested the pitch discrimination thresholds of classically-trained musicians and nonmusicians on a two alternative forced choice task. They found the nonmusicians’ discrimination thresholds to be six times larger than the musicians. Additionally, they gave all participants training on the task and

found that while this did not increase the performance of musicians, nonmusicians improved and performed as well as musicians after eight hours of training. This suggests that pitch discrimination improves with experience, and, at least for discrimination tasks, does not require years of musical experience. Perhaps the ability to discriminate pitch is one that diminishes without sustained practice. During training, the nonmusicians likely received just as much practice interacting with short timescales as the musicians were doing on a continual basis. Speakers of tone languages also regularly engage in pitch discriminations in their daily lives. Pfordresher and Brown (2009) compared a group of nonmusician tone language speakers to a matched sample of nonmusicians and found that the ability to discriminate the interval between two pairs of notes was enhanced in tone language speakers.

The ability to discriminate periodicities at long timescales also differs among individuals, and again it seems to be an experience-based effect. Drake, Penel, & Bigand (2000) investigated the ability of musicians and nonmusicians to tap along with pieces of piano music as an index of discrimination. They tested both expressive (timing elements as played by a real pianist, containing subtle variations called microstructure) and mechanical (no timing microstructure) versions of each piece. Due to the presence or absence of these expressive timing elements, the pulse of each piece would be more or less salient to the listener. They found that synchronization was better with mechanical versions, likely because the stimulus was perfectly periodic. Musicians synchronized their taps more accurately overall and tended to tap at higher referent levels, or at periodicities longer than the pulse corresponding to grouping structure (e.g., 2:1 or 3:1). This suggests that musicians are better able to perceptually group auditory events and attend to multiple periodicities present in a complex stimulus.

The ability to discriminate periodicities at short and long timescales seems to depend on practice. Musicians and tone language speakers frequently make fine pitch discriminations and as a result, both groups have relatively low pitch discrimination thresholds. Musicians also have more experience than nonmusicians with tracking the pulse, leading to the expected practice effect.

Multiple Timescales: Sensorimotor Synchronization

In music, synchronization is essential. Musicians must keep notes in tune, organize note onsets within a phrase, and often coordinate the timing of their playing with other musicians. Coordinating actions over short timescales, such as singing a note in tune with other musicians, is something that improves with practice. There are large individual differences here, with some people readily able to carry a tune while others clearly struggle. Musicians will practice and study for many years to be able to accurately reproduce a desired pitch on command. Pfordresher and Brown's (2009) study of pitch discrimination in tone language speakers also investigated pitch matching via singing, and found that tone language speakers were also more accurate than nonmusicians at reproducing a sequence of tones of different pitch.

Sensorimotor synchronization (SMS)—coordinating perception and action—is an area of research that focuses on synchronization at long timescales, often using tapping tasks (Repp, 2005; Repp & Su, 2013). It has been found that musicians exhibit lower timing variability and smaller asynchronies when finger tapping. Repp (2010) investigated tapping ability in a tapping-continuation paradigm where participants would first synchronize their taps with a pacing signal and continue tapping at the same rate when the pacing signal stopped. He found that musicians would continue tapping at approximately the same rate as the pacing signal, whereas nonmusicians would typically increase their tapping rate. There are even differences within musicians as a group; drummers have

been found to synchronize more accurately than pianists and singers (Krause, Pollok, & Schnitzler, 2010).

If musicians show experience-based behavioural advantages in synchronization tasks, it stands to reason that there should exist neural differences as well. Anatomical changes in auditory cortex have been found in musicians as compared to nonmusicians, with concomitant neurophysiological differences. Schneider et al. (2002) divided their group of participants into professional musicians, amateur musicians, and nonmusicians based on amount of training, performance experience, and a tests of musical aptitude. Using functional magnetic resonance imaging (fMRI), they found that the amount of grey matter in the anteromedial portion of Heschl's gyrus—primary auditory cortex—was significantly greater in professional and amateur musicians than in nonmusicians. Gaser and Schlaug (Gaser & Schlaug, 2003) have found musical experience to correlate with grey matter volume as well, including auditory, motor, somatosensory, and cerebellar regions of the brain. Schneider et al. (2002) also tested magnetoencephalographic (MEG) responses to sinusoidal tones, and found that the evoked response was twice as large for professional musicians than nonmusicians.

The ability to behaviourally synchronize with periodicities at short and long timescales seems to improve with practice. Musicians are often better able to match pitch and carry a tune, presumably because playing music involves sounding tones that fit a tonal and temporal context. Interestingly, these differences seem to be reflected in the brain, since musicians seem to possess more grey matter volume in auditory and sensory-motor regions.

Multiple Timescales: Neural Correlates

Subcortical Synchronization at Short Timescales

When sound enters the cochlea, a neural representation of it passes through the auditory nerve en route to the cortex. Along the way it must pass through various subcortical structures including the cochlear nuclei, superior olivary complex, and the inferior colliculus (IC) (Moore, 2000). Each step in the pathway must carry some representation of the signal, and studies have found that firing patterns in subcortical neurons preserve the oscillatory structure of the stimulating signal at short timescales. Worden and Marsh (1968) made single-cell neural recordings in the cochlear nuclei of cats to tone stimuli and noticed that the response waveform closely matched that of the stimulus. That is, the oscillatory pattern of the tone was present in the brain's EEG response. They called these frequency-following responses (FFR) for their ability to track the periodic characteristics (or frequency) of auditory stimuli. Smith, Marsh, and Brown (1975) extended this work to compare response latency to tone bursts at each of the cochlear nuclei, the medial superior olive, the IC, and the surface of the scalp (Figure 1a). Not surprisingly, they found that response latency increased for structures further down the processing stream, increasing from approximately 2 to 6 ms. Interestingly, the response latency of the IC and the scalp recording were very similar (a difference of only 0.4 ms) leading the authors to posit that the scalp-recorded response has its origins in the IC. Recording scalp responses from humans, they found latencies comparable to those seen in cats (6.5 ms, Figure 1b). Thus, the scalp-recorded FFR in humans is widely believed to originate in the IC. Sohmer, Pratt, and Kinarti (1977) recorded FFRs in a sample of patients with brainstem lesions, and found that periodic elements of the stimulus were not well-represented, strengthening the case for the oscillatory response pattern originating in the IC. Thus, the FFR

offers a unique window into the early neural processing of auditory information.

Aiken and Picton (2008) measured FFRs from normal hearing subjects to two different vowel sounds and found that the spectrum of the FFR closely matched that of the original stimuli. Assuming that energy across the spectrum would be equal in the absence of any stimulus, they found that frequency bins corresponding to f_0 of the stimulus and many of its harmonics had significantly greater amplitude than neighbouring frequencies. Krishnan and Parkinson (2000) played participants pure tones lasting 80 ms that changed in pitch from 400 to 600 Hz, and found that spectral peaks in the FFR closely followed the pitch trajectory of the changing stimulus.

Marmel et al. (2013) investigated the relation between the strength of the FFR to behavioural measures of pitch discrimination. Participants were tested on pure tone audiometry as well as discrimination thresholds for the test frequency (660 Hz) using a three-alternative forced choice procedure. FFRs were collected and they found that a measure of FFR synchronization strength accounted for variance in frequency difference limens beyond that which was accounted for by pure-tone thresholds alone. In the context of the auditory pathway, the subcortical response (FFR) to an auditory stimulus is relayed downstream to cortical areas responsible for higher-level auditory processes. The stronger the FFR, the better one is able to discriminate pitch.

Subcortical Synchronization and Musicianship

For those who interact with pitch discriminations more often, a stronger FFR should be associated with ease of discrimination. Musicians are a good example, since a musician can often tell right away if a piece of music is out of tune. Speakers of tone languages should also show stronger FFR synchronization since certain phonemes are distinguished based on their pitch

characteristics alone. Bidelman, Gandour, and Krishnan (2011) compared FFR responses and discrimination accuracy of English speaking musicians and nonmusicians, and Mandarin speaking nonmusicians to arpeggiated triads (three-tone sequences representing stable positions in major or minor keys). Major and minor triads are very common in Western tonal music and consist of three tones each spaced 3 or 4 semitones apart. Major and minor triads only differ with respect to the f_0 of the tone falling between the highest and lowest tones. Bidelman et al. (2011) used major and minor triads as well as a detuned up version (the middle tone was higher in frequency than in the major triad) and a detuned down version (the middle tone was lower in frequency than in the minor triad) as stimuli. They also asked participants to perform a same-different discrimination task on the three most musically meaningful stimulus pairs (major/minor, major/detuned up, minor/detuned down). They found that spectral peaks in the FFR at the f_0 of the middle tone were significantly larger in musicians and Mandarin-speakers compared to nonmusicians. However, when the Mandarin-speakers were asked to behaviourally discriminate the detuned stimulus pairings they performed just as poorly as nonmusicians. This suggests that experience with discriminating pitch can strengthen synchronization to periodic stimuli in the IC, but there may be multiple domain-specific behavioural consequences of IC synchronization, as evidenced by Mandarin speakers' poor performance on the triad discrimination task despite their relatively strong FFR. The Mandarin speakers' FFR synchronization was on par with musicians, but they lacked the musical experience to convert this into an appropriate behavioural response in the context of a musical task.

Cortical Synchronization at Long Timescales

Evidence for neural synchronization to periodicities at long timescales has also been found. Fujioka et al. (2012a) had participants passively listen to

isochronous sequences at different tempi while they watched a silent movie. Using MEG, they investigated spectral changes in the beta-band—defined as spectral activity in the 13–25 Hz region—in auditory and motor cortices and found a decrease in beta-band power after each tone onset, with the minimum occurring approximately 200 ms following onset, regardless of tempo. However, the time course of the subsequent power increase did change between tempi, such that maximum synchronization occurred simultaneously with the following tone onset. That beta-band power synchronized its oscillation with the periodicity of the stimulus is evidence that the response was not reactionary, but predictive. This pattern of neural activity suggests entrainment rather than passive tracking of the pulse. Much like the FFR, the periodicity of beta-band activity entrains to periodicity present in the stimulus. Activity in the gamma-band (20–60 Hz) has also been found to synchronize with isochronous sequences, and interestingly maintains its synchronized periodicity even when tones are omitted from the sequence (Fujioka, Trainor, Large, & Ross, 2009; Snyder & Large, 2005) or perturbed in time (Zanto, Large, Fuchs, & Kelso, 2005). This potentially points to higher-level predictive functions, such that once cortical circuits have entrained to an oscillation they can endogenously maintain it in the absence of the stimulus. This ability in particular seems well-suited to musicians, since maintaining a beat is critical to playing in time with others.

Using similar metrics to those used in FFR research, Nozaradan et al. (2011) sought to investigate steady-state evoked potentials (SSEP), a phase-locked EEG response that is time-locked to the stimulus. In different conditions they asked their participants to imagine a duple, triple, or no meter (imagining the pulses in groups of two or three or no grouping at all) while listening to an isochronous 2.4 Hz pulse. They found that the spectral pattern in EEG

recordings was determined by the imagined grouping structure. Imagined double groupings elicited a spectral peak at half the pulse frequency (1.2 Hz) and imagined triple groupings elicited a spectral peak at one third (0.6 Hz) and two thirds (1.2 Hz) of the pulse frequency. It seems that the spectrum of evoked EEG activity reflects the perceptual elements of beat and meter.

Cortical Synchronization and Musicianship

Evidence for modulation of cortical synchronization to pulse by musical training is almost non-existent. Fujioka et al. (2012b) investigated beta-band oscillations to isochronous rhythms in stroke patients before and after participating in a music-based movement therapy. They found that beta-band power synchronized with the stimulus, including the consistent beta-band power decrease following each tone reaching its minimum around 200 ms post onset, extending their previous findings in healthy adults (Fujioka, Trainor, Large, & Ross, 2012a). Critically, the beta decrease in the stroke patients was larger after therapy, suggesting a modulation of cortical circuits due to musical training. In a spatial principal components analysis, the contribution of auditory areas seemed to be changed with training. This suggests that similar changes should be present in musicians, since the ability of cortical circuits to track oscillations in at long timescales seems to be modulated with musical training. Since the therapy in Fujioka et al. (2012b) lasted only five weeks, one would expect that many years of musical training and performance would effect more pronounced changes in circuitry to track oscillations at long timescales.

Resonance

The resonance theory of neuronal oscillation (Large, 2008) posits that populations of neurons have the ability to spontaneously synchronize with external stimuli or other populations of neurons. Large (2008) has developed a

mathematical model describing banks of frequency-tuned resonators that oscillate only when certain frequencies are present at their input. An important consequence of the model is that these resonators can also be endogenously activated, suggesting that periodic activity could occur in the absence of any external periodic stimulus. This means that the timing pattern of a rhythm could be generated or maintained without any external synchronization. Another important tenet of the theory is that while resonators respond most strongly to the frequency at which they are tuned, they also respond to related frequencies such as those that are half or twice their tuned frequency. So, it is possible that activity will arise at frequencies that are harmonically related to a stimulus frequency despite not being present in the stimulus.

Evidence for neural resonance can be found in behavioral studies such as the continuation tapping studies described above (Repp, 2010), or studies investigating “the missing fundamental”, where a tone that only contains upper harmonics and no f_0 is still perceived as having pitch, corresponding to f_0 (Fletcher, 1924). Complex rhythms in music rarely contain acoustic events to mark every period of the pulse, yet people can tap along with the pulse of the music. In a follow-up study, Nozaradan, Peretz, and Mouraux (2012) had participants listen to complex rhythms (instead of isochronous ones), and found SSEP spectra that closely matched the spectra of the original rhythms. Despite the fact that many incomplete periodicities were present in the stimulus, populations of neurons were able to synchronize with each frequency and extract a dominant one. Additionally, when participants were asked to tap the perceived beat after the presentation of each rhythm, their tapping coincided with their SSEPs, suggesting that the percepts of pulse and meter are captured in the spectral characteristics of EEG. It seems that the brain is able to track many

loosely delineated periodicities at once, offering further evidence for neural resonance in the perception of periodic stimuli.

Musicianship and Multiple Timescales

It appears that the human auditory system has a remarkable ability to track and synchronize with periodic stimuli, at both short and long timescales. In SMS tasks, performance has been shown to improve with practice, and these improvements are matched with neural changes that allow for better tracking of periodicities. SMS is an important part of playing music, so musicians tend to have an enhanced ability to synchronize (Repp, 2010). Neural synchronization to tones is also enhanced in musicians (Bidelman & Krishnan, 2011; Lee, Skoe, Kraus, & Ashley, 2009), suggesting that the ability of neurons to synchronize with external stimuli contributes to the ability to perceive pitch. Empirical research on individual differences in neural synchronization has focused almost entirely on short timescales, so it remains to be seen whether the enhanced capacity in musicians for neural synchronization extends to long timescales.

There is some evidence that neural synchronization strength in the IC is linked to synchronization in other areas of the brain. Musacchia, Strait, & Kraus (2008) measured EEG responses to the speech sound /*da*/ to investigate differences in subcortical and cortical responses and measures of musicianship. Specifically, they looked at subcortical synchronization to frequencies present in the speech sound, and cortical P1-N1 slope in evoked potentials, a component that has been found to correlate with musical experience (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004). They found that these neural measures correlated well with one another, and also with years of recent musical practice. This indicates that, at least for stimuli of short timescales, subcortical neural synchrony seems to be consistent with cortical evoked responses. The relation

between musicianship and synchronization at long timescales, however, has not been investigated.

Thus, the current proposal seeks to investigate whether musicians will show stronger neural synchronization with auditory stimuli at short timescales as well as long timescales. Similar to the FFR extracted from subcortical activity, stronger synchronization at the cortical level should manifest as an SSEP, showing more spectral energy at the pulse frequency of an isochronous sequence.

Taken a step further, it can be hypothesized that individuals with better synchronization ability at the subcortical level will also have better synchronization ability at the cortical level. Since synchronization to temporal structure is an ability that seems to be improved by musical training (Bidelman & Krishnan, 2011), it seems reasonable to posit that the enhancement would improve proportionately across different timescales. As such, the strength of subcortical synchronization at short timescales (tones) should correlate with the strength of cortical synchronization at long timescales (pulse).

Additionally, neural measures of synchronization should be able to predict behavioural measures of musical experience. To investigate this, a regression analysis will be performed on measures of musical experience, using subcortical and cortical indices of neural synchronization as predictors.

Methods

Participants

Thirty-three people with a wide range of musicianship were recruited from the Ryerson University student participant pool and from the community at large, and received either course credit or an honourarium, respectively. For group analyses, musicians and nonmusicians were categorized based on criteria used in previous studies investigating musicianship (Bidelman et al., 2011; Lee et al., 2009). Musicians (6) had at least 10 years of experience playing their instrument (mean \pm SD; 19.5 ± 6.3) beginning at or before the age of 11 (7.2 ± 3.1) and were active musicians at the time of testing. Nonmusicians (4) had at most 2 years of experience playing an instrument (1.1 ± 1.0) and had not been musically active for at least 7 years (17.2 ± 7.8). Those not meeting either of these criteria (7) were categorized as amateur musicians and their data were excluded from most group analyses to maximize group differences and to model the analyses on previous work¹. All participants reported having normal hearing. Participants who spoke tone languages (4) were excluded from all analyses. Other participants were excluded due to artifacts (6; described in detail in the analysis section) and various technical difficulties related to the running the experiment (6)².

¹ It was originally planned to only compare two groups (musicians and nonmusicians, as in Bidelman et al., 2011), but given a large amount of variability in musicianship in the sample it was decided to have three groups (including amateur musicians, as in Lee et al., 2009).

² For three participants there was electrical noise in the FFR recording due to an un-grounded computer power cable on the recording computer. One participant had electrical noise in his/her SSEP recording due to a damaged electrode lead. One participant heard a stimulus with wavering volume due to an unknown computer problem. One participant did not hear the stimulus at all (experimenter error).

Stimuli and Procedure

Behavioural tasks. *Pitch discrimination.* Stimuli for the pitch discrimination task were implemented with the psychoacoustics toolbox for MATLAB (Grassi & Soranzo, 2009). The psychoacoustics toolbox uses an adaptive maximum likelihood procedure to estimate auditory thresholds (Grassi & Soranzo, 2009), and contains procedures for evaluating intensity thresholds, gap detection thresholds, and pitch discrimination thresholds, among others. For the pitch discrimination task, tones consisted of 4 equal-amplitude harmonics and were 250 ms long. A tone with a f_0 of 220 Hz was always taken as the standard, and the target f_0 ranged between 220 and 260 Hz. On each trial, participants heard three tones, two of which were standard and the other the target, and were instructed to respond which tone—the first, second, or third—was the target by entering their response on a computer keyboard.

Pulse perception. The Beat Alignment Test (BAT; Iversen & Patel, 2008) measures pulse perception in music in the absence of overt synchronization. It contains musical excerpts from a variety of popular styles (rock, jazz, pop orchestral) with varying saliency of pulse. On each trial, the excerpt plays for 5 seconds before a sequence of isochronous beeps are imposed on the music, at which point the participant is instructed to respond whether the beeps coincide with the beat of the music. The BAT was implemented as a computer-based task using custom MATLAB scripts.

EEG task. EEG stimuli consisted of complex tones with a f_0 of 220 Hz and its first 5 partials (440, 660, 880, 1100, and 1320 Hz), all of equal amplitude (See Figure 2; Bidelman & Krishnan, 2009). Each had a duration of 80 ms with a Hanning window amplitude envelope and was exported with a sampling rate of 20000 Hz. Tones were presented in an isochronous sequence at a rate of 2.5 Hz (an inter-onset-interval (IOI) of 400 ms). Thus, the stimulus train was created to

have periodic properties at both short and long timescales, such that synchronization from both subcortical and cortical sources could be simultaneously recorded from the same stimulus. An isochronous train of tones achieves this, since each tone oscillates at a certain frequency (short timescale), and the rate of presentation oscillates at a different frequency (long timescale). Presentation consisted of 5136 tones (lasting approximately 33 minutes). Responses from the first 16 tones were discarded to a) avoid movement artifact since the experimenter left the room during the presentation of these tones, and b) to maximize pulse perception in participants, since multiple cycles of a rhythm are required before pulse perception occurs (Chapin et al., 2010). A loudness accent was put on the first of every group of four tones to induce a percept of meter. Since imagining a binary or ternary beat on an unaccented train of tones will give rise to neural activation at different frequencies (Nozaradan et al., 2011), this was done to minimize variability in modes of listening across participants.

General procedure. Upon arrival at the lab, participants completed a consent form and were seated comfortably in a sound-attenuated booth in front of a computer screen and keyboard. They were fitted with magnetically shielded insert earphones (Etymotic ER-3A); all stimuli were presented monaurally in the right ear (Bidelman et al., 2011). Participants first completed 36 trials of the BAT and two 15-trial blocks of the pitch discrimination task. They were then fitted with EEG electrodes for recording both subcortical (BIOPAC Systems, Inc.) and cortical (BioSemi Instrumentation) neural activity. Each EEG response was recorded on a separate computer. Participants then listened to the isochronous train; they were instructed to view a silent movie and to disregard to the tones. This method of passive listening is widely used in human research on subcortical synchronization (Skoe & Kraus, 2010), with some studies even

allowing participants to fall asleep during recording (Bidelman et al., 2011). Despite previous work in pulse perception using complex rhythms showing that attention is necessary to obtain a stable pulse percept (Chapin et al., 2010), other work using isochronous stimulus trains like the one in this study have shown neural synchronization during passive listening (Fujioka et al., 2009; Fujioka, Trainor, Large, & Ross, 2012a).

EEG stimuli were presented at 70 dBA using Presentation (Neurobehavioral Systems) on a different computer than the one used to play the movie. Lastly, participants completed two questionnaires: a general music questionnaire assessing their history of playing music, listening habits, and current musical activities, and the Absorption in Music Scale (AIMS; Sandstrom & Russo, 2013), which measures an individual's willingness to allow music to draw them into an emotional experience.

EEG recording and processing. *Subcortical data.* As per the manual for the recording hardware (BIOPAC Systems, Inc.), brainstem activity was recorded from the ipsilateral (active electrode) and contralateral (reference) earlobes with a forehead electrode placed just below Fpz (ground). There is no common method of electrode placement in FFR research, with other studies including Cz and the mastoid bone for the placement of active or reference electrodes. The arrangement in the current study (i.e., earlobes and forehead) was advantageous because the 64-channel electrode cap prevented access to Cz or the mastoids. Data were recorded at a sampling rate of 20000 Hz using AcqKnowledge 4.1 software; the auditory stimulus was also recorded on a separate channel to facilitate segmenting the data³. Data were hardware filtered online between 1 and 3000 Hz, exported to MATLAB files, and analyzed using custom MATLAB

³ The audio signal from the computer was split using a hardware splitter to both present it to the participant and record it for segmenting purposes, using it as a trigger. This ensures that the trigger is accurately timed to the stimulus.

scripts. The entire response train was bandpass filtered between 150 and 1500 Hz using a 3rd-order Butterworth filter. Epochs were identified using the stimulus recording, and were segmented into windows starting 120 ms before the midpoint of the stimulus and extending 220 ms. Trials exceeding $\pm 35 \mu\text{V}$ were discarded (Bidelman & Krishnan, 2011).

One issue with recording FFR responses is the cochlear microphonic (CM), where the resonance of the cochlea to the auditory stimulus can contaminate the response of the brainstem. The time course of the FFR and the CM can be clearly delineated since the CM is nearly coincident with the stimulus whereas the FFR occurs with a 6-10 ms delay (Skoe & Kraus, 2010; Smith et al., 1975). Thus the CM can be identified by comparing the timing of stimulus and response peaks. Additionally, the phase of CM responses remains constant with respect to the stimulus, but the phase of the FFR does not. As a result, a common method of dealing with the CM is to create two versions of the stimulus in opposite polarity (i.e., multiple the stimulus waveform by -1). If an equal number of CM responses to each polarity are averaged together the result will be zero, thus eliminating the artifact from the recording⁴.

The tones in the current study were presented in each polarity in blocks of 64 tones, and an equal number of positive- and negative-polarity responses were averaged together. Each polarity was identified on the basis of order, such that once all stimulus onsets were identified and the first 16 tones discarded, the next 64 tones were positive polarity, the next 64 tones were negative polarity, and so

⁴ Three participants had a cochlear microphonic that was louder than the FFR as discerned by the fact that response peaks were not delayed by >6 ms (Skoe & Kraus, 2010). These participants were removed from all analyses.

on ⁵. This meant discarding some good responses, but including more of a single polarity in the average would reintroduce the cochlear microphonic.

Since the FFR occurs with a delay of 6-10 ms (Smith et al., 1975), the target window began 10 ms after stimulus onset (to focus on the periodic portion of the response) and extended to 10 ms after stimulus offset, with the baseline period extending 80 ms prior to stimulus onset (Skoe & Kraus, 2010). Grand averaged waveforms for musicians and nonmusicians are shown in Figure 3a.

For each participant, a fast-Fourier transform (FFT) was applied to the resulting average waveforms for both the baseline and target regions in 5 Hz bins (Lee et al., 2009). To eliminate the noise floor, the amplitude spectrum of the baseline region was subtracted from the amplitude spectrum of the target region (Skoe & Kraus, 2010). The result was a single spectrum for each participant, with the energy at target frequency bins being taken as the measure of synchronization. Grand averaged spectra for musicians and nonmusicians are shown in Figure 3b.

Cortical data. Cortical responses were recorded with a 64-channel electrode cap (BioSemi ActiveTwo System) and digitized at a sampling rate of 512 Hz, allowing for a frequency response from DC to 256 Hz. Raw cortical EEG data were preprocessed using EEGLAB functions. First, the data were highpass filtered at 0.1 Hz to eliminate slow drifts, and noisy channels and data were removed using the `clean_rawdata` plugin in EEGLAB (Delorme & Makeig, 2004). All channels were referenced to the average reference and subjected to an independent components analysis (Makeig, Bell, Jung, & Sejnowski, 1996).

⁵ For three participants the recording of the FFR stimulus had a signal-to-noise ratio that was too small for all stimulus onsets to be found. This was due to improperly-matching impedance levels between the audio jack and the earphones (earphones were plugged into the speaker jack). Thus it could not be confirmed that the resulting average would not contain an equal number of opposite polarity stimuli, risking the presence of the cochlear microphonic. These participants were removed from all analyses.

Artifactual components including eye blinks and eye movements were identified using the ADJUST plugin (Mognon, Jovicich, Bruzzone, & Buiatti, 2011). The contribution of these components was removed from the timeseries data to repair these artifacts, and the result was epoched into 25-second segments corresponding to 64 tones each (ignoring the first 16 tones). For each participant and electrode, segments were averaged together in the time domain to remove activity that was not phase-locked to the stimulus (evoked activity), and a FFT with a frequency bin size of 0.0312 Hz was calculated (Nozaradan et al., 2011). Since each segment was immediately preceded by another segment, there was not an opportunity to use a pre-stimulus baseline region to remove the contribution of unrelated spectral noise, as has been done in prior FFR studies (Skoe & Kraus, 2010). Instead, the contribution of the noise floor was removed by subtracting, from each frequency bin, the mean of the closest two frequency bins that lie at least two bins away from the current bin (Nozaradan et al., 2011). For example, if bin 5 is the current bin, then the mean of bins 1, 2, 8, and 9 would be subtracted from bin 5. This procedure assumes that in the absence of any stimulus, the energy at each frequency bin should be similar to the mean of neighbouring bins. The total energy at pulse-related frequencies for each participant and electrode was taken as the sum of three bins centered on the target frequency, and these values were averaged across all electrodes (Nozaradan et al., 2011). Grand averaged spectra for musicians and nonmusicians are shown in Figure 4.

Analysis and Results

Group Differences. Behavioural measures. To maximize group differences and in line with previous studies (Bidelman et al., 2011; Lee et al., 2009), amateur musicians were left out of the group analyses. Musicians had more years of experience playing music than did nonmusicians ($t(8)=5.63, p<0.001$)

and also spent more hours per week playing ($t(8)=3.888, p=0.005$) and listening ($t(8)=2.396, p=0.043$) to music. They also had significantly higher AIMS scores ($t(8)=3.842, p=0.005$) indicating that they are more willing to become emotionally absorbed into a piece of music. Musicians had lower pitch discrimination thresholds (3.5 ± 2.0 Hz compared to 7.6 ± 6.4 Hz for nonmusicians) and higher BAT accuracy scores (0.84 ± 0.14 compared to 0.66 ± 0.17 for nonmusicians) than nonmusicians, but independent samples t -tests did not reach significance ($t(8)=-1.499, p=0.172$, and $t(8)=1.817, p=0.107$, respectively).

Neural synchronization to f_0 and the pulse. A 6 x 3 mixed ANOVA was run on the subcortical data to investigate if harmonics were differentially represented in the FFR across groups, despite having equal amplitude in the stimulus. The within-subjects factor in this analysis was Frequency (6 levels) and the between-subjects factor was Group (3 levels). The ANOVA found a significant main effect of Frequency ($F(5,11)=11.001, p<0.001$) and a nonsignificant effect of Group ($F(2,14)=0.306, p=0.741$). There was also a significant interaction of Frequency x Group ($F(10,70)=2.538, p=0.011$). Analysis of Figure 6a suggests that synchronization was skewed to f_0 and especially so in musicians relative to nonmusicians.

The spectral peaks in neural activity at the frequency of f_0 and the pulse were of particular interest as indices of neural synchronization (Lee et al., 2009). Differences between musicians and nonmusicians in subcortical synchronization to f_0 of the tone (220 Hz) approached significance ($t(8)=2.010, p=0.079$, Figure 5a), while cortical synchronization to the pulse frequency did not significantly differ between groups ($t(8)=0.884, p=0.402$, Figure 5b). Subcortical peaks at the harmonics of the tone and cortical peaks at pulse-

related frequencies also showed no significant differences (See Table 1 and Figures 6a and 6b).

Correlational Analyses. For correlational and regression analyses, the group of amateur musicians were included as well as musicians and nonmusicians (Lee et al., 2009). Neural indices of synchronization at short and long timescales were marginally correlated ($r=0.45$, $p=0.068$), but neither correlated with years of musical experience ($r=0.09$, $p=0.73$ for short timescales and $r=0.06$, $p=0.82$ for long timescales) despite what has been previously reported (Lee et al., 2009). However, synchronization at short timescales did correlate with the amount of time participants spent playing music ($r=0.55$, $p=0.021$, Figure 7a), and synchronization at long timescales was marginally correlated with the amount of time spent listening to music ($r=0.46$, $p=0.067$, Figure 7b).

Regression analyses. To further investigate the relation between musical engagement and neural synchronization, a regression analysis was run to see if the number of hours each participant spends playing or listening to music could be predicted by a linear combination of their spectral FFR and SSEP peaks for synchronization at short and long timescales. The regression explained approximately 31% of the variance and approached significance ($R^2=0.314$, $p=0.071$; see Figure 8).

Discussion

The current thesis investigated differences in neural synchronization between musicians and nonmusicians. It was found that while musicians did show larger spectral peaks at the frequency of the stimulus at both short and long timescales, these differences were only marginally significant. They did not correlate with years of musical experience as predicted, however they did correlate with the number of hours per week spent playing or listening to music. Previous research indicates that musical experience relates to the strength of subcortical synchronization (Bidelman et al., 2011; Lee et al., 2009), but the current study suggests a more nuanced explanation. The amount of time currently spent engaged with music might be better predictor of neural synchronization than total years of musical experience. Additionally, the strength of neural synchronization at short and long timescales together can predict the amount of time an individual spends engaged in musical activities. This kind of coherence in synchronization across timescales and levels of musicianship is a novel finding and suggests that synchronization abilities in different parts of the brain may be improved concurrently with musical engagement. Since music contains periodicities at different timescales, it makes sense that an ability to faithfully represent them neurally would be necessary for understanding music as a listener or interacting with it as a player.

Musicianship and Neural Synchronization

The ANOVA revealed a main effect of Frequency, suggesting that neural synchronization is stronger to certain frequencies compared to others. Since the stimulus frequencies were all of equal amplitude, this is an interesting finding and suggests that subcortical synchronization is focused on certain aspects of the stimulus. In particular, synchronization to f_0 was greater than for other frequencies. This is perhaps not surprising since the frequency of f_0 is what

defines the pitch of a tone, and this pitch defines how the tone musically interacts with other tones in melody and harmony, for example. The significant interaction between Frequency and Group adds to this interpretation, as musicians had a subcortical synchronization profile across frequencies that was more skewed towards f_0 , presumably due to their experience interacting with pitch, melody and harmony.

Subcortical synchronization to tones was found to correlate with playing music while cortical synchronization to the pulse was marginally correlated with listening to music. This is curious considering that periodicities at both short and long timescales would be present while both playing and listening to music, so it would be expected that synchronization strength would increase at both timescales. There are many possibilities for this pattern of results, but one explanation is the following. Outside of music playing and listening, there is much opportunity in daily life to synchronize movements with events at long timescales, such as synchronizing steps when walking with another person, turn-taking during conversation, or coordinating movements while playing team sports. Opportunities to synchronize with short timescales are relatively less frequent, and seem almost exclusive to music-making (singing) and tone language speaking. Given this, we would expect that a task only practiced by musicians, namely playing an instrument, would correlate with synchronization at short timescales, and that all participants might synchronize reasonably well at long timescales. While listening to music, one might inadvertently synchronize their movements with the beat of the music such as walking on the beat or bobbing their head. Music provides an explicit opportunity to synchronize with the beat, so we might expect that those who listen to more music are more practiced at synchronizing at long timescales.

While playing, an individual is engaged with the music and is actively synchronizing their movements to play the correct notes at the correct time⁶. In contrast, the listener is only perceiving the auditory display, so it might even be expected that playing would correlate better than listening at both timescales. Taking a closer look at the data in the current study, the number of hours spent playing music among the sample of musicians was mostly made up of solo playing as opposed to playing in a group setting with other musicians. The FFR peak correlated better with solo playing ($r=0.55$, $p=0.022$) than playing in a group ($r=0.44$, $p=0.075$). It could be argued that solo playing involves less pressure on the timing elements of the piece, and more emphasis on the pitch of notes. For example, when practicing by one's self a musician might stop right away to correct a pitch error, despite effectively halting the metric structure. This puts more emphasis on the correct pitches, leading to more practice synchronizing at short versus long timescales. In contrast, when playing in a group, pitch errors must be ignored in favour of maintaining the beat, thus putting more emphasis on synchronization at long timescales. A listener might not be interested in actively synchronizing with the music, but oscillations at long timescales can be seen readily in daily life—the interval between steps when walking, for example. Listeners might inadvertently synchronize or more readily comprehend the metrical structure of a piece compared to its harmonic structure. The pattern of results presented here suggests that synchronization at different timescales improve independently of one another. Since music contains both pitches and rhythms, synchronization ability tends to improve concurrently at both timescales.

⁶ Which note and time are the “correct” ones is usually up for debate, and there might be many different “correct” ones depending on the style of music and the performer. Here, “correct” is taken to mean notes that generally fit the harmonic and metrical structure of the piece.

These experience-based changes are perhaps more specific than predicted, and musicians of different types might synchronize differently depending on how much they engage with oscillations at various timescales in their daily life. For example, singers and violin players might show enhanced synchronization at short timescales since they must tune each note that they sing or play, while drummers might show enhanced synchronization at long timescales since they are most concerned with the timing of note onsets and synchronizing to the beat. There is some precedence for this; Martens (2011) had participants tap along with musical excerpts and found that different people preferred to tap at different metrical levels. He classified three different groups based on preferred metrical level: deep tappers who prefer tapping at higher metrical levels (grouping more events with each tap), surface tappers who prefer tapping at lower metrical levels (grouping fewer events with each tap), and variable tappers who tended to tap at diverse metrical levels. Interestingly, deep tappers tended to have musical training in low range instruments (e.g., bass), conducting, composition, and dance, whereas surface and variable tappers tended to have training in high range instruments (e.g., violin, guitar), piano, voice, and percussion. This suggests that there are different aspects of music in which one can have expertise, and it stands to reason that these would manifest in neural as well as behavioural changes. Future work involving musicians should endeavour to more precisely quantify exclusion criteria. The seemingly subtle differences between different musicians current and past experience with music can come to bear on the results of studies of this kind.

We know that spending time playing music will increase grey matter volume in auditory areas such as Heschl's gyrus (Schneider et al., 2002), but there has not been much research about how these benefits might deteriorate over time once music playing has stopped. The current study suggests that

stopping or reducing the amount of time engaged with music will effect neural changes; neural indices of synchronization to music correlated with the current, and not past, amount of musical experience. Neurally, music is a complex sensorimotor act that requires the precise integration of a variety of brain regions, so it is conceivable that these circuits would require ongoing maintenance to retain their synchronization ability. Of course, musicians who stop playing their instrument lose the adeptness that they may have once had. Given the close coupling between neural and behavioural measures of musicianship, it makes sense that a reduction in behavioural ability would coincide with some deterioration of related neural representations.

Synchronization Across Multiple Timescales

The current study suggests coherence in neural synchronization ability across cortical levels and musical experience, but whether or not these operate as separate or coordinated mechanisms is unknown. As mentioned above, the differential pattern of correlations of playing and listening to music suggests that synchronization at different timescales might improve independently of one another. However, since sensory signals must pass through subcortical regions prior to arriving at cortical ones (Moore, 2000), the quality of the cortical signal is likely entirely dependent on the quality of the subcortical signal. This suggests that synchronization improvements should occur first in subcortical areas and later on in cortical areas. It also seems possible that subcortical mechanisms could poorly represent oscillatory activity at short timescales, while still accurately relaying oscillatory activity at long timescales (i.e., note onsets) further down the processing stream.

FFR synchronization to speech sounds has been shown to correlate with cortical evoked potentials (Bidelman, Weiss, Moreno, & Alain, 2014; Musacchia et al., 2008), so there seems to be some coherence between subcortical and

cortical responses to stimuli at short timescales. Tierney and Kraus (2013) measured IC synchronization strength to the speech syllable /*da*/ as well as behavioural measures of tapping variability while synchronizing with a metronome. They found IC synchronization to be related to tapping synchronization accuracy, leading the authors to suggest that links between the IC and the cerebellum lead to coordinated improvements in synchronization. It seems that improvements in neural synchrony at one timescale lead to improvements at other timescales as well, suggesting that the mechanism of synchronization might be the same regardless of timescale or cortical level. Exactly how synchronization at multiple timescales and cortical levels might be related to one another is a topic that requires further study, but the current study finds that while musical engagement seems to improve synchrony at both timescales, the mechanisms underlying each likely operate independently.

Complex Periodicities

The current study focused on periodicities that did not change: the pitch of tones and the pulse rate were constant throughout. How neural synchronization responses behave with dynamic stimuli is still an open question, though the ability of the FFR to track changes in pitch has been studied. For example, the FFR has been found to closely track the pitch contour of a tone that was changed dynamically between 400 and 600 Hz (Krishnan & Parkinson, 2000). There is also evidence for robust subcortical representation of speech sounds that change their pitch such as those in tone languages, for both tone language speakers as well as musicians (Krishnan, Xu, Gandour, & Cariani, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007). If neural synchronization behaves the same way at longer timescales, then a cortical representation of tempo changes in an isochronous rhythm should be just as strong. Future work should investigate the ability of SSEPs to track these changing periodicities.

The pulse frequency in a piece of music is rarely explicitly marked with an auditory event at each pulse, so the listener must infer the pulse frequency based on the relative locations of events in time. In complex rhythms many periodicities are present at once, and none of them has all onsets marked. Surely neural resonance is involved (Large & Snyder, 2009), where synchronization at one frequency enhances perception at related frequencies. How we come to know where the pulse frequency is within a complex rhythm could be elucidated by studying the nature of neural synchronization. Nozaradan (2012) had participants listen to different non-isochronous sequences and found many spectral SSEP peaks at many of the periodicities that were present in the stimuli. The stimuli contained many different and incompletely marked periodicities; that the brain is still able to represent them is evidence of a kind of neural resonance.

The notion of groove—defined as the desire to move part of one’s body in time with music (Janata, Tomic, & Haberman, 2012)—might come to bear on the strength of cortical synchronization at long timescales as well. In a study investigating the acoustic properties of high and low groove music, Madison et al. (2011) found that the salience of the beat correlated with participants’ groove ratings regardless of genre. It is presumably easier to behaviourally synchronize with a more salient beat, suggesting that high groove music would show enhanced SSEPs at the pulse frequency. Since there is variability among individuals in the ability to feel the pulse of music, it could be argued that those who are better able to extract the pulse—musicians, for example—would more often feel the “groove.” The potential interaction between pulse salience and ability to extract the pulse seems well-suited to investigating how neural oscillations might give rise to certain perceptions.

Higher-Level Synchronization Processes

Short and long timescales involve stimuli that are vastly different perceptually, yet they can be considered to both lie on a single continuum spanning both timescales. Although tones and rhythms are thought of as two very different—perhaps opposite—elements of music, in terms of neural synchronization their differences can be characterized better as one of degree, and not one of kind. That neurons at both cortical levels are able to synchronize their firing with stimuli spanning both timescales suggests that other higher- and lower-level brain functions might operate in the same way. Buzsaki (2006) has suggested that networks of neurons can synchronize with external stimuli as well as with other neurons, enabling a form of communication between brain areas. Based on synaptic transmission times, synchronization among neurons could indicate the size of the network, with smaller networks showing synchronization at short timescales and larger networks at long timescales. This kind of hierarchical organization loosely mirrors the organization of periodicities in music, with tones (regions of local oscillation) being organized into the temporally larger concept of meter. Perhaps this is no surprise since brains are ultimately the entity that creates music in the first place, and as such music is well-poised to offer a window into brain function more generally. As Buzsaki has commented, “if self-generated brain dynamics have a link to the spectral composition of speech and music, one might expect that the same dynamics would influence a plethora of other behaviours” (2004, p. 123). The specific mechanism of higher-level synchronization includes stimuli that are potentially too complex to study, but the basic principles of oscillation, synchronization, and resonance would still hold. Emotions such as empathy could be the result of neurons synchronizing with various aspects of the perceptual display that are integrated together to create an internal

representation. Neurons are oscillatory in order to perceive and interact with the external world.

Limitations

The experimental setup employed here is unprecedented and presented considerable challenges due to the simultaneous recording of both subcortical and 64-channel cortical EEG activity. Musacchia, Strait, and Kraus (2008) measured activity from both subcortical and cortical levels using a single active electrode placed at Cz, using highpass and lowpass filters to extract subcortical and cortical data, respectively. Since the current study aimed to replicate methods of recording SSEPs (Nozaradan et al., 2011)—which have been recorded with 64-channel EEG—as well as FFR, two separate recording systems were used. The grounding electrode for the FFR was placed on the forehead, resulting in more electrical noise in frontal electrodes (Fpz, Fp1, Fp2, AFz, AF3, AF4, AF7, AF8). The logistics of physically fitting the FFR electrodes and insert earphones amongst the cortical EEG equipment was also a challenge that had to be overcome. Despite these challenges, both subcortical and cortical data were of good quality.

The current study brings into question the usefulness of using group analyses in studies investigating musicianship. There are surely experience-based effects of musicianship on many perceptual, behavioural, and neural measures, but there is so much variability within musicians and nonmusicians that it becomes difficult to attribute differences to a single metric like musical training or experience. For example, it could be the case that listeners attend more to short or long timescales based on preference. Given that there is variability in the preferred metrical level with which to tap along (Martens, 2011), there might also be variability in the preferred timescale to which to attend. Individuals who interact more frequently with one level of periodicity

(such as nonmusician tone language speakers or dancers) might reveal a more lopsided pattern of synchronization across timescales. Future work should take care to identify appropriate aspects of musicianship and musical experience that are most relevant to the study being conducted.

Conclusion

The current thesis suggests that the ability of neurons to synchronize their firing with external periodic stimuli is something that is modulated by experience and that likely must be maintained with practice. Neural activity did not correlate with years of musical experience as expected, but rather with the amount of time the individual is currently engaged playing or listening to music. The pattern of results also hints towards separate neural pathways for encoding stimuli at different timescales, as playing and listening to music may rely more heavily on synchronization at short or long timescales, respectively. Music perception and production is a seemingly complex task that involves tracking and interacting with a dynamic display of periodicities at multiple timescales. Neural oscillations share many properties with the perception and production of music, and can offer insight into the mechanisms of other complex human behaviours.

Tables

Independent Samples t-tests Between Musicians and Nonmusicians

| <u>Behavioural Measures</u> | <u>t</u> | <u>df</u> | <u>p</u> | <u>Effect Size (r)</u> |
|------------------------------------|----------|-----------|-----------|------------------------|
| Years of Musical Experience | 5.628 | 8 | 0.0005*** | 0.894 |
| Hours/Week Playing Music | 3.888 | 8 | 0.005** | 0.809 |
| Hours/Week Listening to Music | 2.396 | 8 | 0.043* | 0.646 |
| Pitch Discrimination Threshold | -1.499 | 8 | 0.172 | 0.468 |
| Beat Alignment Test Accuracy | 1.817 | 8 | 0.107 | 0.540 |
| Absorption in Music Scale | 3.842 | 8 | 0.005** | 0.805 |
| | | | | |
| <u>Subcortical Synchronization</u> | <u>t</u> | <u>df</u> | <u>p</u> | <u>Effect Size (r)</u> |
| 220 Hz (f_0 of tone) | 2.010 | 8 | 0.079 | 0.579 |
| 440 Hz | 0.187 | 8 | 0.856 | 0.066 |
| 660 Hz | -0.652 | 8 | 0.533 | 0.225 |
| 880 Hz | -0.294 | 8 | 0.776 | 0.103 |
| 1100 Hz | -0.760 | 8 | 0.469 | 0.259 |
| 1320 Hz | -0.475 | 8 | 0.647 | 0.166 |
| | | | | |
| <u>Cortical Synchronization</u> | <u>t</u> | <u>df</u> | <u>p</u> | <u>Effect Size (r)</u> |
| 0.625 Hz | 0.242 | 8 | 0.815 | 0.085 |
| 1.25 Hz | 1.139 | 8 | 0.288 | 0.374 |
| 2.5 Hz (pulse frequency) | 0.884 | 8 | 0.402 | 0.298 |
| 5 Hz | -0.873 | 8 | 0.408 | 0.295 |
| 7.5 Hz | 0.305 | 8 | 0.768 | 0.107 |
| 10 Hz | -0.312 | 8 | 0.763 | 0.110 |

Table 1. Independent samples *t*-tests on spectral peaks (subcortical and cortical) between musicians and nonmusicians. Asterisks indicate *t*-tests significant at the 0.05 (*), 0.01 (**), or 0.001 (***) levels.

MUSICIANSHIP AND NEURAL SYNCHRONIZATION

Correlations Between Behaviour and Neural Synchronization

Correlations with FFR Peak at 220 Hz

| <u>Behavioural Measure</u> | <u><i>r</i></u> | <u><i>p</i></u> | <u><i>Power</i></u> |
|--------------------------------|-----------------|-----------------|---------------------|
| Years of Musical Experience | 0.089 | 0.733 | 0.0531 |
| Hours/Week Playing Music | 0.555 | 0.021* | 0.672 |
| Hours/Week Listening to Music | 0.360 | 0.156 | 0.306 |
| Pitch Discrimination Threshold | 0.091 | 0.727 | 0.0539 |
| Beat Alignment Test Accuracy | 0.314 | 0.219 | 0.240 |
| Absorption in Music Scale | 0.336 | 0.187 | 0.270 |

Correlations with SSEP Peak at 2.5 Hz

| <u>Behavioural Measure</u> | <u><i>r</i></u> | <u><i>p</i></u> | <u><i>Power</i></u> |
|--------------------------------|-----------------|-----------------|---------------------|
| Years of Musical Experience | -0.061 | 0.817 | 0.0423 |
| Hours/Week Playing Music | 0.232 | 0.369 | 0.147 |
| Hours/Week Listening to Music | 0.454 | 0.067 | 0.470 |
| Pitch Discrimination Threshold | -0.021 | 0.937 | 0.0301 |
| Beat Alignment Test Accuracy | -0.198 | 0.447 | 0.118 |
| Absorption in Music Scale | 0.254 | 0.325 | 0.170 |

Table 2. Correlations between neural measures of synchronization (subcortical and cortical) and behavioural measures of musicianship. Asterisks indicate *t*-tests significant at the 0.05 (*), 0.01 (**), or 0.001 (***) levels.

Figures

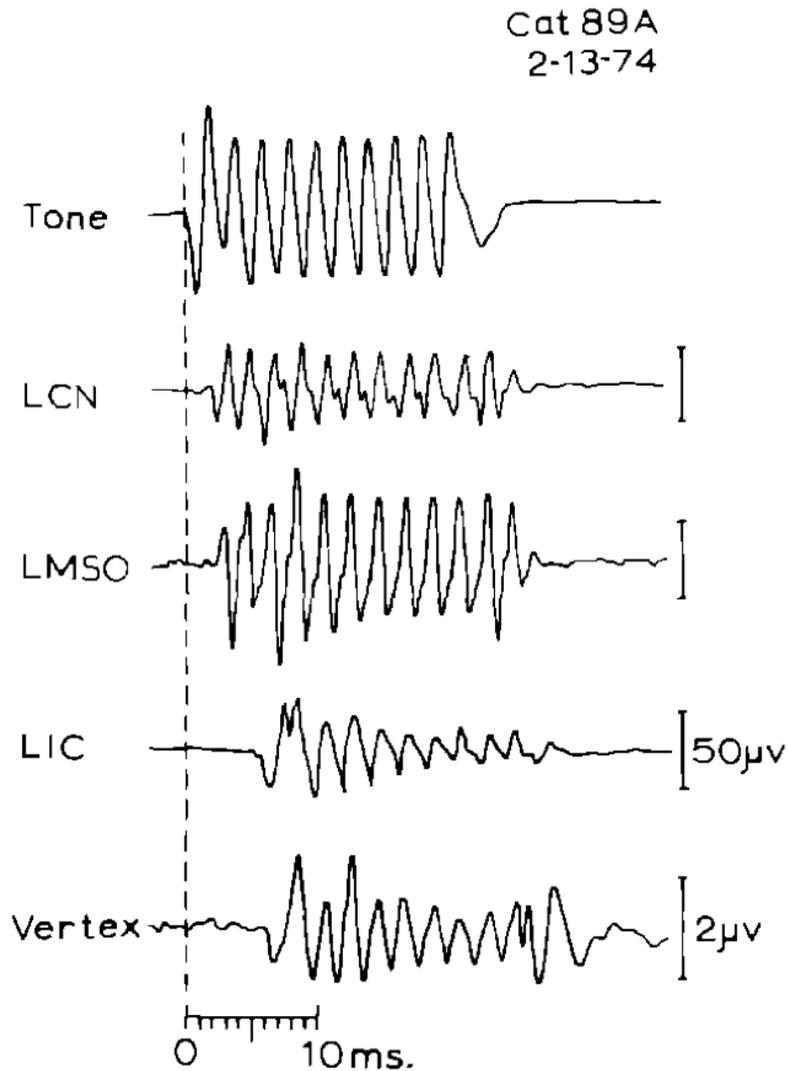


Figure 1a. Single-cell recordings from various auditory structures in the cat to a tone stimulus. The responses mimic the oscillatory structure of the stimulus, and response latency increases for structures further down the processing stream. LCN: left cochlear nucleus; LMSO: left medial superior olive; LIC: left inferior colliculus. Figure from Smith et al. (1975).

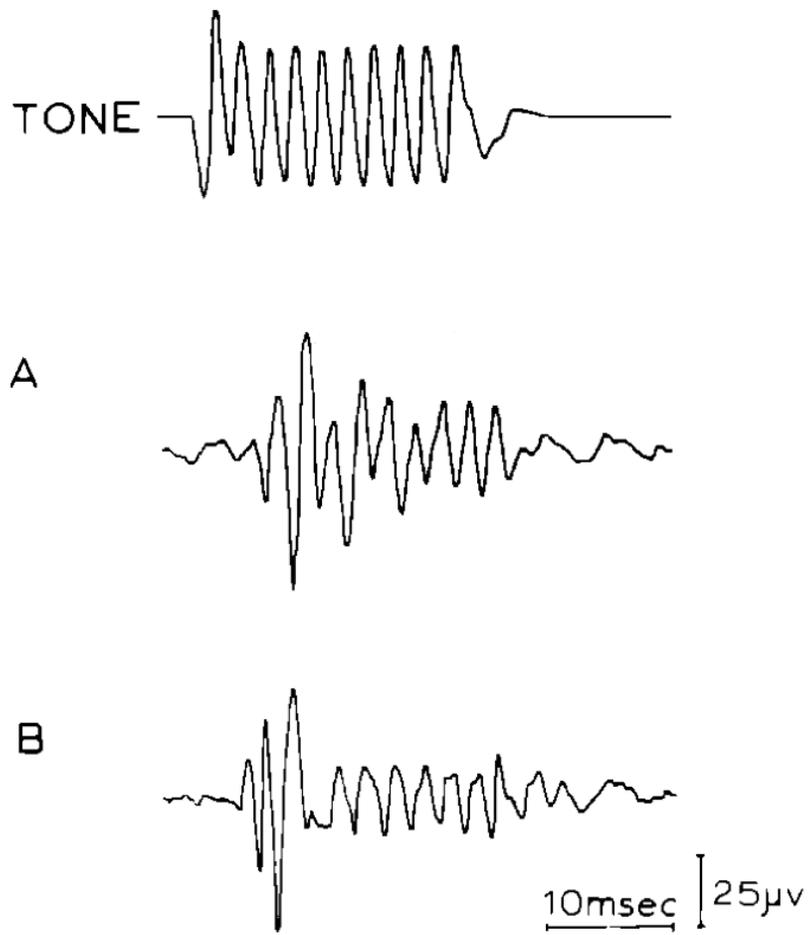


Figure 1b. Recordings from the vertex of the scalp to a tone stimulus from a cat (A) and a human (B). The latency of each response is similar, suggesting that the responses have similar origins. Single-cell recordings in cats suggest this response originates in the IC, leading to the conclusion that the human response also originates in the IC. Figure from Smith et al. (1975).

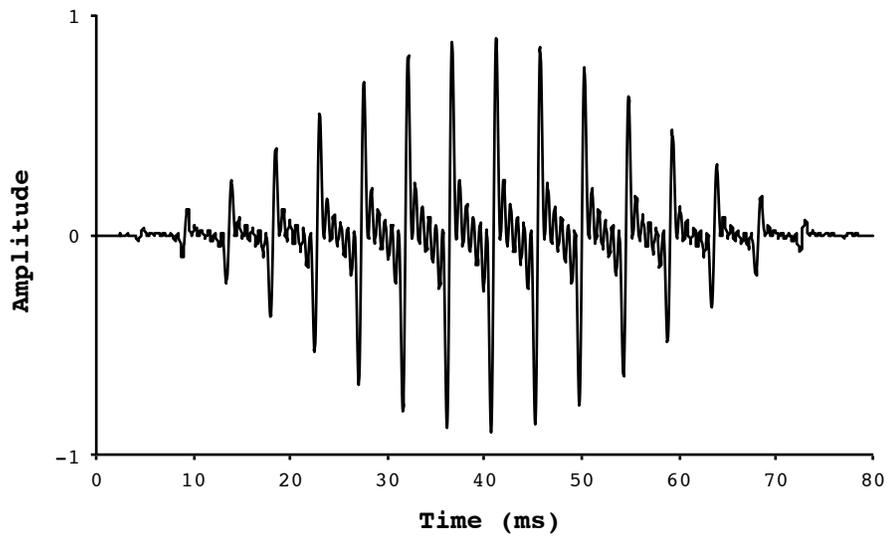


Figure 2a. Waveform of the tone stimulus. The isochronous stimulus train consisted of 5136 of these tones presented at a rate of 2.5 Hz.

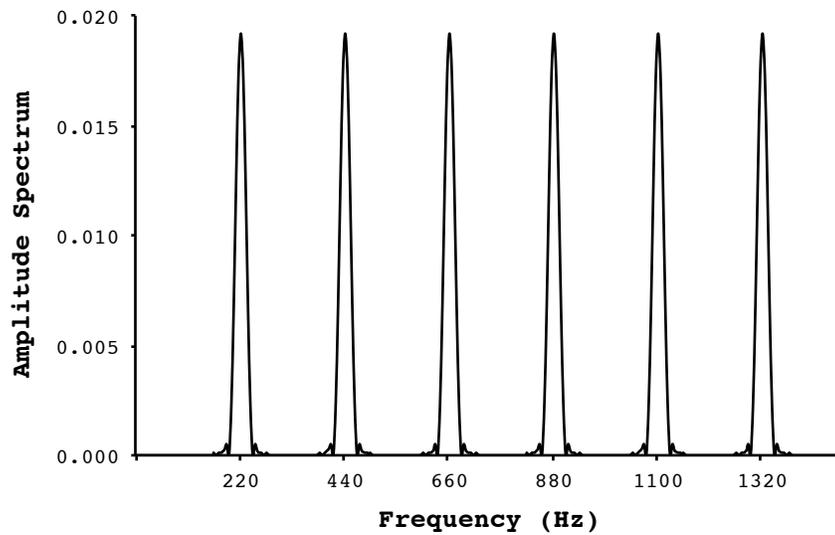


Figure 2b. Amplitude spectrum of the stimulus. Each tone had a fundamental frequency of 220 Hz and the first 5 harmonics of equal amplitude.

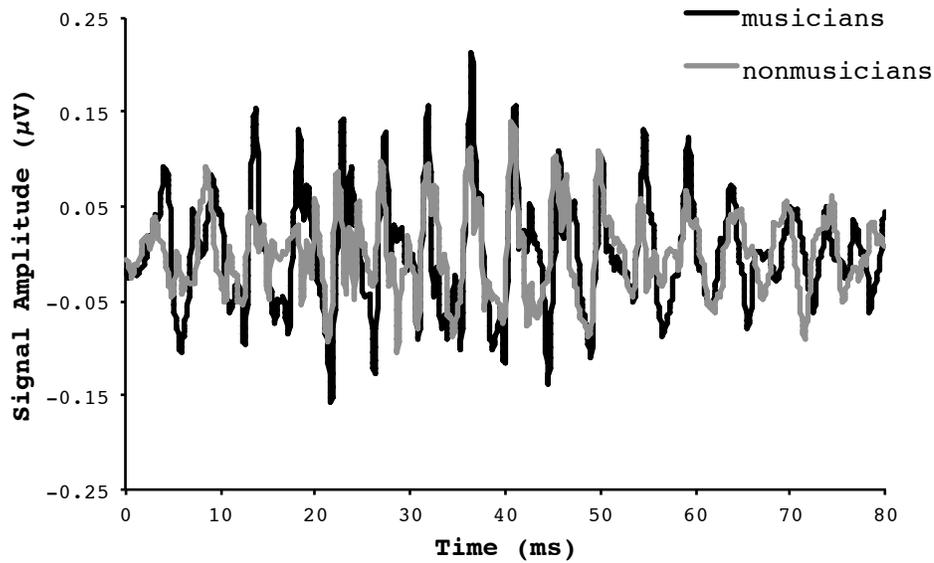


Figure 3a. Frequency following responses for musicians and nonmusicians. The amplitude difference between peaks and troughs is larger for musicians than nonmusicians.

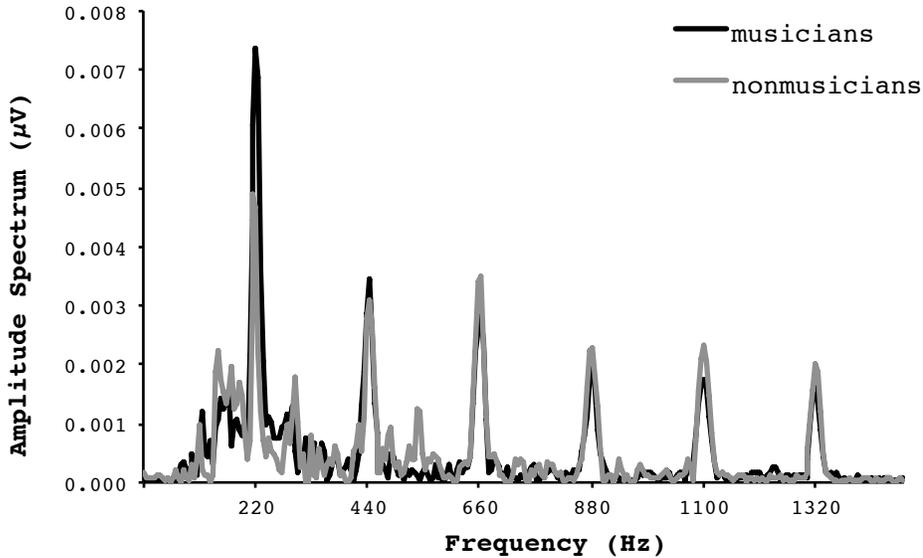


Figure 3b. FFR amplitude spectra for musicians and nonmusicians. The larger peak for musicians at 220 Hz indicates stronger neural synchronization to the f_0 of the tone. Synchronization is seen at other harmonics as well.

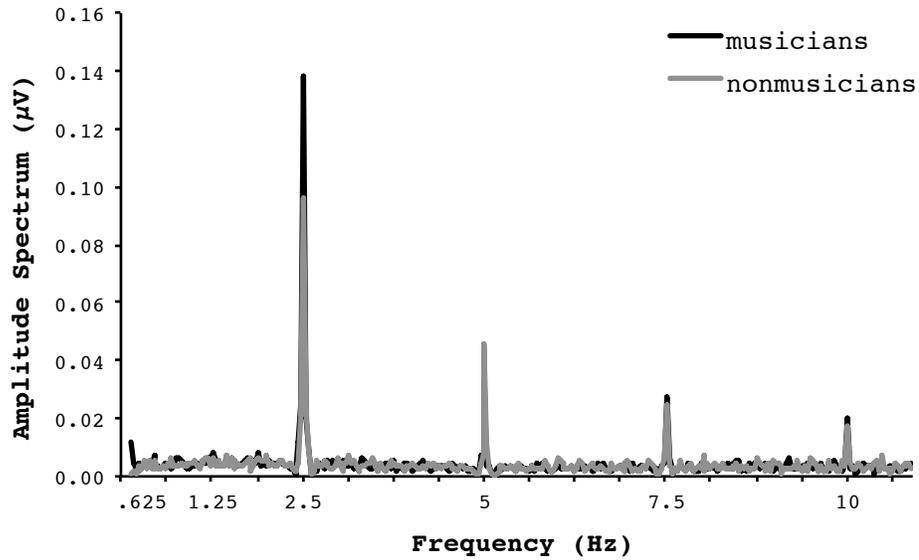


Figure 4. SSEP amplitude spectra for musicians and nonmusicians. The larger peak for musicians at 2.5 Hz indicates stronger neural synchronization at the pulse frequency. Neural resonance is seen at frequencies related to the pulse.

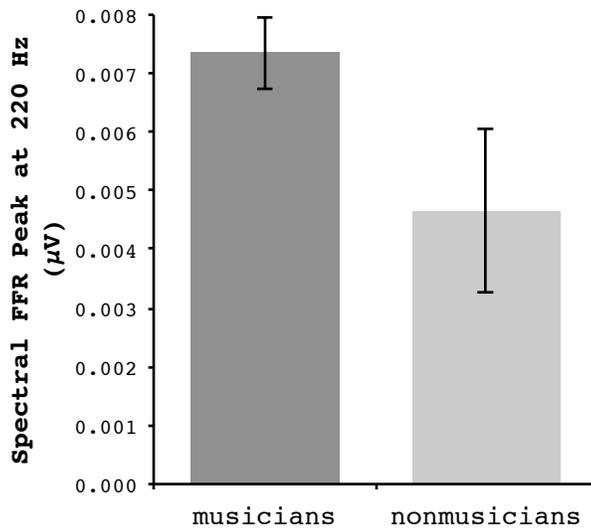


Figure 5a. Subcortical synchronization at f_0 of the tone stimulus (220 Hz). Marginal group differences were found ($t(8)=2.010$, $p=0.079$). Error bars indicate standard error.

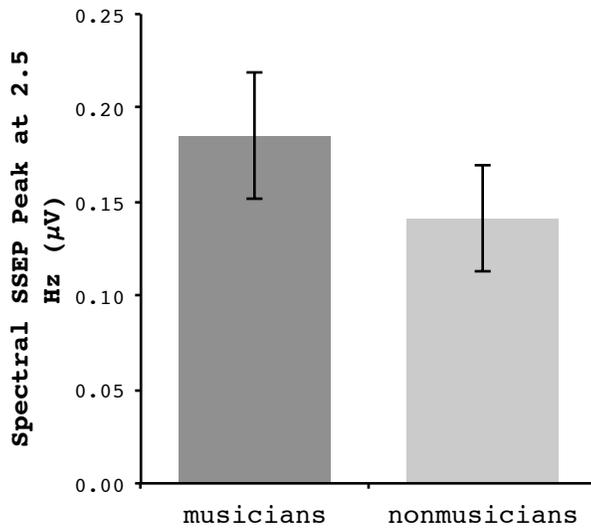


Figure 5b. Cortical synchronization at the frequency of the pulse (2.5 Hz). No significant group differences were found ($t(8)=0.884$, $p=0.402$). Error bars indicate standard error.

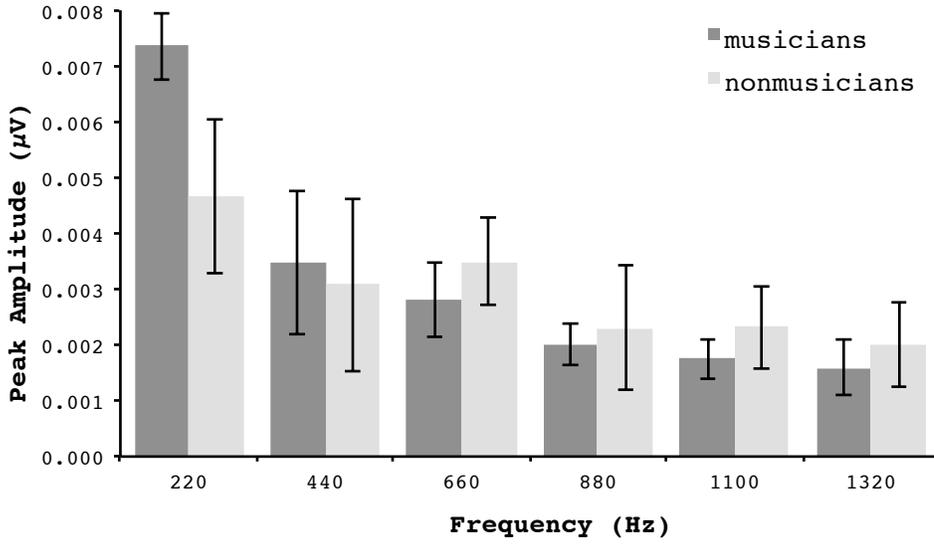


Figure 6a. Subcortical synchronization to harmonics of the tone. Error bars indicate standard error.

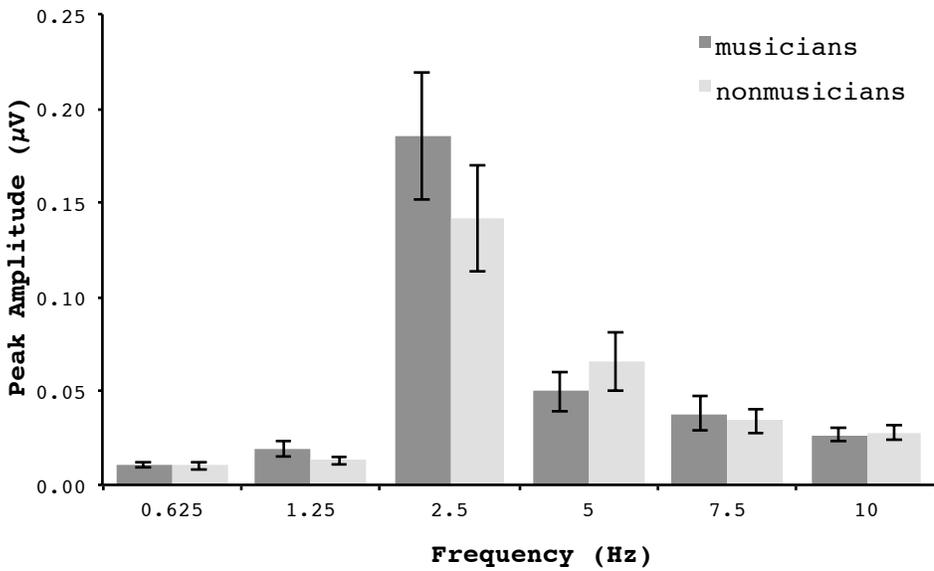


Figure 6b. Cortical synchronization to the pulse and related frequencies. Error bars indicated standard error.

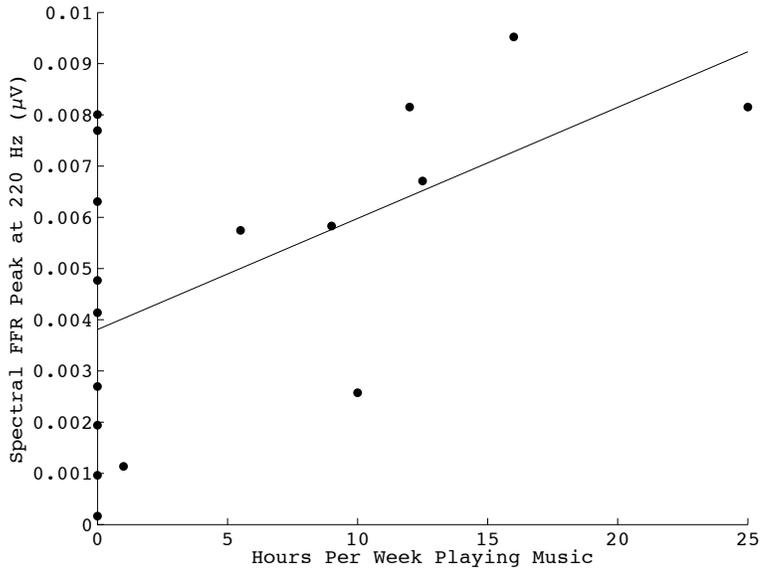


Figure 7a. Spectral FFR peaks correlate with hours playing music ($r=0.56$, $p=0.021$). Subcortical synchronization to f_0 of a tone is related to the amount of time an individual spends playing music.

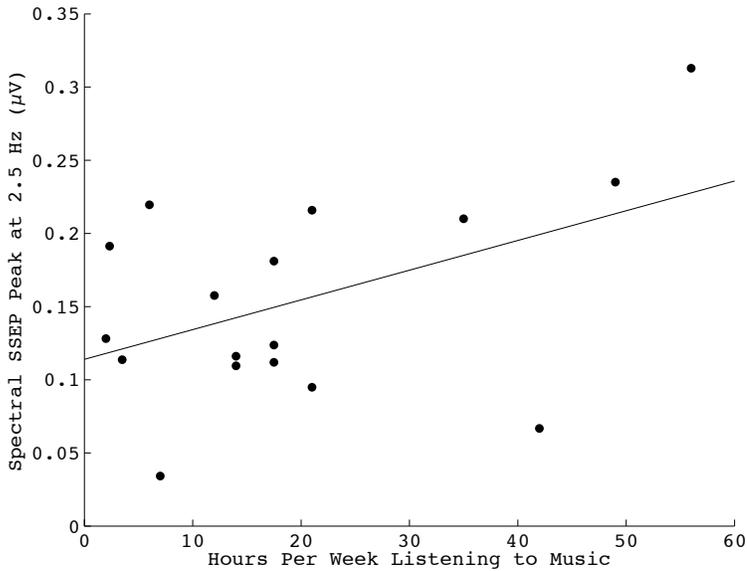


Figure 7b. Spectral SSEP peaks are marginally correlated with hours spent listening to music ($r=0.46$, $p=0.067$). Cortical synchronization to a pulse may be related to the amount of time an individual spends listening to music.

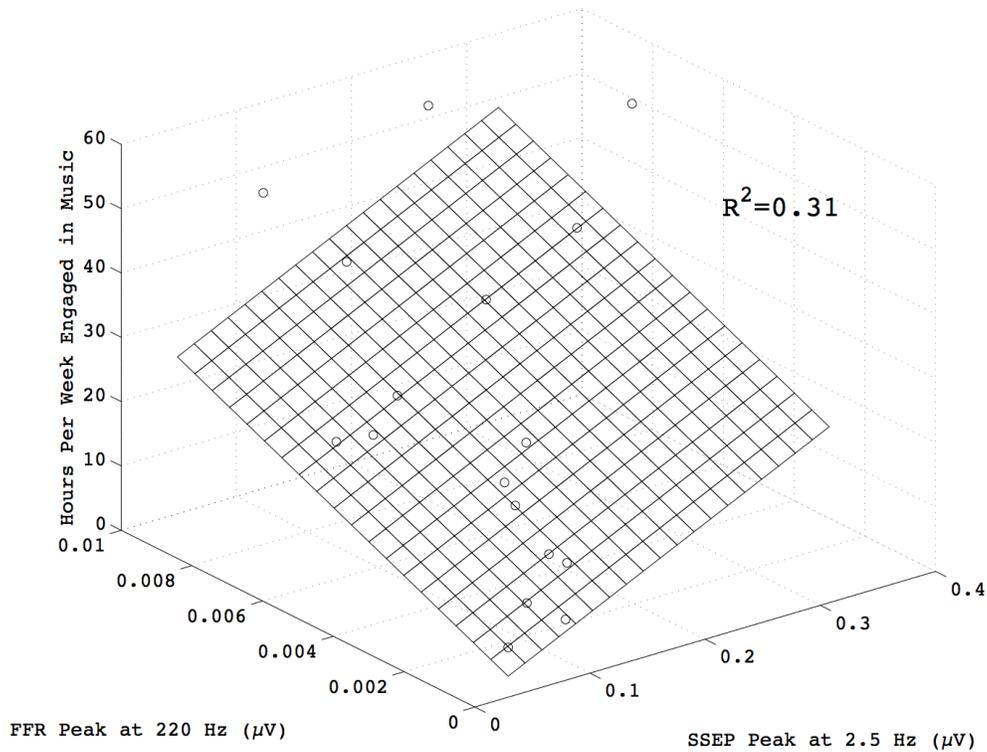


Figure 8. Neural synchronization at short and long timescales predicts musical engagement. A linear regression model using the FFR peak at 220 Hz and the SSEP peak at 2.5 Hz as predictors can successfully predict the number of hours participants spend playing or listening to music ($R^2=0.314$, $p=0.071$).

References

- Aiken, S. J., & Picton, T. W. (2008). Hearing Research. *Hearing Research*, 245(1-2), 35–47. doi:10.1016/j.heares.2008.08.004
- Bidelman, G. M., & Krishnan, A. (2009). Neural Correlates of Consonance, Dissonance, and the Hierarchy of Musical Pitch in the Human Brainstem. *Journal of Neuroscience*, 29(42), 13165–13171. doi:10.1523/JNEUROSCI.3900-09.2009
- Bidelman, G. M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *NeuroReport*, 22(5), 212–216. doi:10.1097/WNR.obo13e328344a689
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Brain and Cognition. *Brain and Cognition*, 77(1), 1–10. doi:10.1016/j.bandc.2011.07.006
- Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience*, n/a–n/a. doi:10.1111/ejn.12627
- Buzsaki, G. (2006). *Rhythms of the Brain*. New York, NY: Oxford University Press.
- Buzsaki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, 304, 1926–1929.
- Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S. J. A., Steinberg, F., & Large, E. W. (2010). Neural Responses to Complex Auditory Rhythms: The Role of Attending. *Frontiers in Psychology*, 1. doi:10.3389/fpsyg.2010.00224
- Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognitive Brain Research*, 80, B1–B10.

- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009
- Drake, C., Penel, A., & Bigand, E. (2000). Tapping in Time With Mechanically and Expressively Performed Music. *Music Perception*, 18, 1–23.
- Fletcher, H. (1924). The Physical Criterion for Determining the Pitch of a Musical Tone. *Physical Review*, 23(3), 427–437.
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and Gamma Rhythms in Human Auditory Cortex during Musical Beat Processing. *Annals of the New York Academy of Sciences*, 1169, 89–92. doi:10.1111/j.1749-6632.2009.04779.x
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012a). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *Journal of Neuroscience*, 32(5), 1791–1802. doi:10.1523/JNEUROSCI.4107-11.2012
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical Training Enhances Automatic Encoding of Melodic Contour and Interval Structure. *Journal of Cognitive Neuroscience*, 16(6), 1010–1021.
- Fujioka, T., Ween, J. E., Jamali, S., Stuss, D. T., & Ross, B. (2012b). Changes in neuromagnetic beta-band oscillation after music-supported stroke rehabilitation. *Annals of the New York Academy of Sciences*, 1252, 294–304. doi:10.1111/j.1749-6632.2011.06436.x
- Gaser, C., & Schlaug, G. (2003). Brain Structures Differ between Musicians and Non-Musicians. *The Journal of Neuroscience*, 23(27), 9240–9245.

- Grassi, M., & Soranzo, A. (2009). MLP: A MATLAB toolbox for rapid and reliable auditory threshold estimation. *Behavior Research Methods*, *41*(1), 20–28. doi:10.3758/BRM.41.1.20
- Hevner, K. (1935). The Affective Character of the Major and Minor Modes in Music. *The American Journal of Psychology*, *47*, 103–118.
- Iversen, J. R., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population (pp. 465–468). Presented at the Proceedings of the 10th International Conference on Music Perception and Cognition, Sapporo, Japan.
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*(1), 54–75. doi:10.1037/a0024208
- Krause, V., Pollok, B., & Schnitzler, A. (2010). Perception in action: The impact of sensory information on sensorimotor synchronization in musicians and non-musicians. *Acta Psychologica*, *133*(1), 28–37. doi:10.1016/j.actpsy.2009.08.003
- Krishnan, A., & Parkinson, J. (2000). Human Frequency-Following Response: Representation of Tonal Sweeps. *Audiology & Neuro-Otology*, *5*, 312–321.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariansi, P. A. (2004). Human frequency-following response: representation of pitch contours in Chinese tones. *Hearing Research*, *189*, 1–12. doi:10.1016/S0378-5955(03)00402-7
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys. *Psychological Review*, *89*(4), 334–368.
- Large, E. W. (2008). Resonating to Musical Rhythm: Theory and Experiment. In S. Grondin, *Psychology of Time* (pp. 189–231). Bingley, UK: Emerald Group Publishing Limited.

- Large, E. W., & Snyder, J. S. (2009). Pulse and Meter as Neural Resonance. *Annals of the New York Academy of Sciences*, 1169, 46–57. doi:10.1111/j.1749-6632.2009.04550.x
- Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective Subcortical Enhancement of Musical Intervals in Musicians. *Journal of Neuroscience*, 29(18), 5832–5840. doi:10.1523/JNEUROSCI.6133-08.2009
- Madison, G., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modeling the Tendency for Music to Induce Movement in Humans: First Correlations With Low-Level Audio Descriptors Across Music Genres. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1578–1594. doi:10.1037/a0024323
- Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analysis of Electroencephalographic Data. In D. Touretzky, M. Mozer, & M. Hasselmo, *Advances in Neural Information Processing Systems 8* (pp. 145–151). Cambridge, MA: MIT Press.
- Marmel, F., Linley, D., Carlyon, R. P., Gockel, H. E., Hopkins, K., & Plack, C. J. (2013). Subcortical Neural Synchrony and Absolute Thresholds Predict Frequency Discrimination Independently. *Journal of the Association for Research in Otolaryngology*, 14(5), 757–766. doi:10.1007/s10162-013-0402-3
- Martens, P. A. (2011). The Ambiguous Tactus: Tempo, Subdivision Benefit, And Three Listener Strategies. *Music Perception: an Interdisciplinary Journal*, 28(5), 433–448. doi:10.1525/mp.2011.28.5.433
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, 219(1-2), 36–47. doi:10.1016/j.heares.2006.05.004
- Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and

- temporal features. *Psychophysiology*, 48(2), 229–240. doi:10.1111/j.1469-8986.2010.01061.x
- Moore, J. K. (2000). Organization of the Human Superior Olivary Complex. *Microscopy Research and Technique*, 51, 403–412.
- 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(1-2), 34–42. doi:10.1016/j.heares.2008.04.013
- Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm. *Journal of Neuroscience*, 32(49), 17572–17581. doi:10.1523/JNEUROSCI.3203-12.2012
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the Neuronal Entrainment to Beat and Meter. *Journal of Neuroscience*, 31(28), 10234–10240. doi:10.1523/JNEUROSCI.0411-11.2011
- Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics*, 71(6), 1385–1398. doi:10.3758/APP.71.6.1385
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6), 969–992. doi:10.3758/BF03206433
- Repp, B. H. (2010). Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Human Movement Science*, 29(2), 200–213. doi:10.1016/j.humov.2009.08.002
- Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*. doi:10.3758/s13423-012-0371-2

- Sandstrom, G. M., & Russo, F. A. (2013). Absorption in music: Development of a scale to identify individuals with strong emotional responses to music. *Psychology of Music, 41*(2), 216–228. doi:10.1177/0305735611422508
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience, 5*, 688–694. doi:10.1038/nn871
- Skoe, E., & Kraus, N. (2010). Auditory Brain Stem Response to Complex Sounds: A Tutorial. *Ear & Hearing, 31*(3), 302–324. doi:10.1097/AUD.0b013e3181c8b272
- Smith, J. C., T, M. J., & Brown, W. S. (1975). Far-Field Recorded Frequency-Following Responses: Evidence for the Locus of Brainstem Sources. *Electroencephalography and Clinical Neurophysiology, 39*, 465–472.
- Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research, 24*, 117–126. doi:10.1016/j.cogbrainres.2004.12.014
- Sohmer, H., Pratt, H., & Kinarti, R. (1977). Sources of Frequency Following Responses (FFR) in Man. *Electroencephalography and Clinical Neurophysiology, 42*, 656–664.
- Tierney, A., & Kraus, N. (2013). The Ability to Move to a Beat Is Linked to the Consistency of Neural Responses to Sound. *Journal of Neuroscience, 33*(38), 14981–14988. doi:10.1523/JNEUROSCI.0612-13.2013
- Webster, G. D., & Weir, C. G. (2005). Emotional Responses to Music: Interactive Effects of Mode, Texture, and Tempo. *Motivation and Emotion, 29*(1), 19–39. doi:10.1007/s11031-005-4414-0

- 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*. doi:10.1038/nn1872
- Worden, F. G., & Marsh, J. T. (1968). Frequency-Following (Microphonic-Like) Neural Responses Evoked by Sound. *Electroencephalography and Clinical Neurophysiology*, 25, 42–52.
- Zanto, T. P., Large, E. W., Fuchs, A., & Kelso, J. A. S. (2005). Gamma-Band Responses to Perturbed Auditory Sequences: Evidence for Synchronization of Perceptual Processes. *Music Perception*, 22, 535–552. doi:10.1111/j.1467-9280.1997.tb00536.x