6-2009

Music and Language: Exploring an Artificial Music Grammar

Erica R. Knowles
Seton Hall University

Recommended Citation

Knowles, Erica R., "Music and Language: Exploring an Artificial Music Grammar" (2009). *Seton Hall University Dissertations and Theses (ETDs)*. 2383.
https://scholarship.shu.edu/dissertations/2383

MUSIC AND LANGUAGE:
EXPLORING AN ARTIFICIAL MUSIC GRAMMAR

By
Erica R. Knowles

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Experimental Psychology with a concentration in Behavioral Neuroscience.
Department of Psychology
Seton Hall University

June,2009

Approved By:

Marianne E. Lloyd

Dr. Marianne E. Lloyd, Faculty Mentor

Kelly M. Goedert

Dr. Kelly Goedert, Committee Member

Janine C. Buckner

Dr. Janine Buckner,
Director of Graduate Studies and Committee Member

For Mom and Dad.

Thank you for always letting me follow my own path
no matter how many turns it has taken.

Acknowledgements

I would like to sincerely thank Dr. Marianne Lloyd for her guidance and support throughout my two years at Seton Hall. I am extremely grateful to her for allowing me to explore and develop a research project that was truly my own despite the fact that it was outside her realm of expertise. I could not have asked for a better mentor. I would also like to thank the other members of my committee, Dr. Kelly Goedert and Dr. Janine Buckner, for their support and assistance in completing this project. In these three individuals I have found exemplars of the kind of researcher and teacher I hope to become.

I am extremely grateful towards my fellow graduate students in whom I have found great friendships and to the members of the Memory Lab for their assistance with this project. I would also like to express my gratitude to Dr. Robert Sleve, who was kind enough to allow me to use his sentence stimuli for this project.

My family has supported me throughout my life and without their love I would not be the person I am today. I am extremely grateful for their encouragement and their continued interest in my research. I would also like to thank my friends who have supplied me with distractions and kept me grounded throughout this process.

Finally, I would like to thank my loving husband Ian. He has been incredibly patient and has constantly reminded me that the most important things in life are often outside of school.

Table of Contents

List of Figures

List of Tables

Abstract

Research regarding the brain mechanisms that underlie music and language processing supports two main interpretations: domain-specificity and domain-generality. Evidence from neuropsychology literature, specifically from amusia research, supports domain-specific mechanisms (Peretz & Coltheart, 2003) but recent neuroimaging and behavioral evidence supports overlapping mechanisms, especially for syntax processing (Patel, 2008). The present study used an artificial music grammar in order to test participants' ability to learn a new music grammar as well as to observe a possible interaction between music and language syntax processing. Although participants were able to learn the artificial music grammar, a language task was not affected by errors in the new grammar as has been found with Western music-syntax errors (Sleve, Rosenberg, & Patel, 2009). Future research should consider extending exposure to the artificial grammar to allow for better learning in order for errors in the new grammar to affect the processing of language syntax.

Introduction

Two major organized sound systems are found throughout human culture, language and music. While these systems have many obvious differences, they also share some major similarities. One of these similarities is the use of rule-based combinatorial sequences. Both music and language are created through the combination of discrete elements, notes and chords in music and words in language. These elements form sequences based on hierarchical principles and are therefore both syntactic systems. New evidence from neuroimaging and behavioral research points to a possible overlap in the processing, including syntactic processing, of music and language (Patel, 2003, 2007). This research challenges the long held belief that language and music processing relies on domain specific neural systems (Peretz & Coltheart, 2003).

Grammatical Structures in Music and Language

Syntax (a.k.a. grammar) is the organization of discrete elements (e.g. words in language and pitches in music) through a set of principles which govern how these elements are combined to form sequences (e.g. sentences in language and melodies in music: Jackendoff, 2002). Both language and musical sequences are not created through the random combination of these basic elements but rather their organization is governed by a set of combinatorial principles. These principles act to govern the formation of words, phrases, and sentences in language, and of chords, progressions, and keys in music (Patel, 2003).

The majority of research on music uses Western tonal music as stimuli and an introduction to this idiom may prove helpful for the following discussion. There is a large body of research on the structure and perception of Western tonal music and it is one of the most widely practiced musical systems of the developed world (Patel, 2003). Another reason for the use of Western tonal music is that there are enough structural norms in how the music is composed to enable researchers to study the grammar of music (Patel et al., 1998). This musical idiom consists of a finite set of pitches (a.k.a. tones), which can be represented by those found on a piano keyboard, with note names A through G. These 12 tones (A–A#–B–C–C#–D–D#-E–F–G–G#–A) are then combined into subsets of seven tones, the scales (Tillman, Janata, & Bharucha, 2003). These scales consist of seven notes and are the building blocks of musical pieces (Peretz, 1993). These scales also define musical keys which are important elements of musical syntax (Sieve, Rosenberg, & Patel, 2009).

Within a key, scale tones (the pitches that make up the scale) are organized around a central pitch known as the tonic (Figure 1). This central tone defines the key, that is, in the key of G major the pitch G is the tonic. The other scale tones within the key are arranged into a hierarchy of importance or stability (Krumhansl, 1990). The fifth and the third tones of the scale are considered to be the most clearly related to the tonic and therefore the most stable. The remaining scale tones are considered to be less related to the tonic. Tones from outside a specific key are the least related and often sound like 'foreign' or 'alien' tones within the context of that key (Peretz & Coltheart, 2003). An easy way to think of this is that the seven tones that make up the scale (and therefore key) are the in-key tones and therefore those remaining five tones are the out-of-key tones (Tillman, Janata, & Bharucha, 2003). For example, the key of C major contains the pitches C–D–E–F–G–A–B. These tones are the in-key tones and C#, D#, F#, G#, and A#

are the out-of-key tones (See Figure 1). It is the way in which these tones are arranged to form an "event hierarchy" that defines the musical syntax of a piece (Patel, 2007).

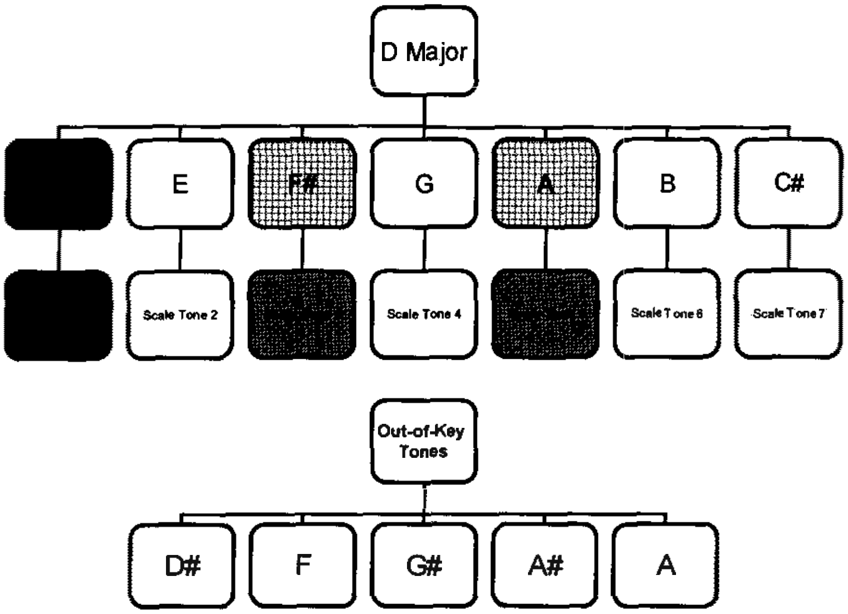

Figure 1. Western Tonal Scale. The top of the figure represents the key of D major. Listed are the scale tones of that key. The bottom of the figure shows the tones that would out-of-key in the context of D major.

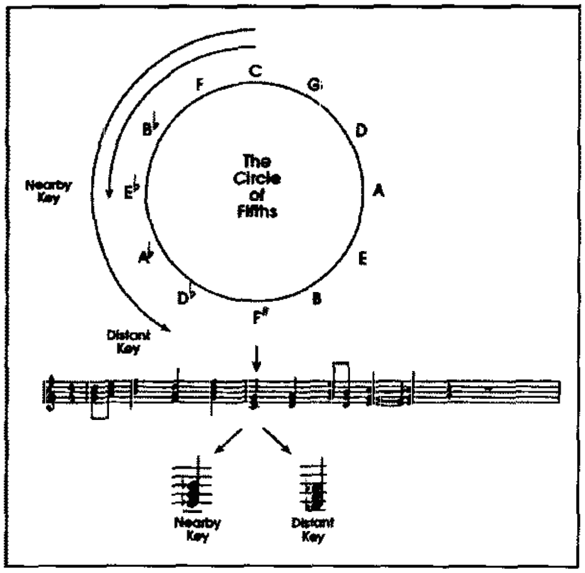

Figure 2 (Patel, 1998). The top portion of the figure shows the circle of fifths, a way in which the Western tonal keys are organized. A nearby key is one that is close in proximity to the target chord on the circle of fifths and therefore contains tones that may overlap with the target key. The distant keys are much further away from the target key on the circle of fifths. Distant keys would share very few tones with the target key. The bottom on the figure shows a chord from a nearby key and a chord from a distant key. As you can see, the nearby chord shares a tone with the target chord while the distant chord does not.

Musical keys also have a systematic organization among themselves (Figure 2). Those keys that share multiple tones are heard by listeners as being more closely related to one another. The more distant keys are from one another, the perceived relatedness decreases. If a tone is structurally unexpected it results in a musical syntactic error and difficulty in processing occurs; in that, the listener's expectancies have been violated and therefore comprehension is disrupted because of the need to integrate the syntactically unexpected event within the preceding context. This processing difficulty is very similar to what occurs with a syntactic error in language (Sieve et al., 2009).

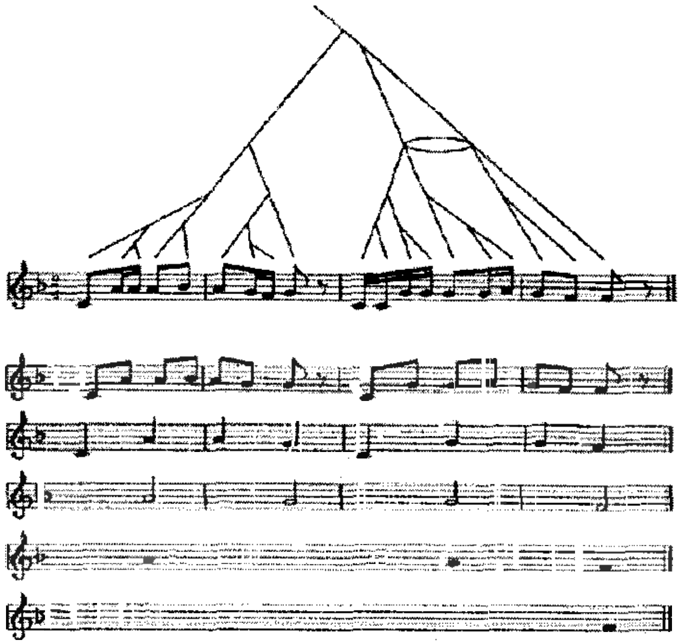

Figure 3 (Patel, 2007). A time-span reduction of the first two phrases of the children's song "Hush little baby" (tree notation from Lerdahl & Jakendoff, 1983). Shorter branches terminate on less important pitches, while longer branches terminate on more important pitches. The lower staves show the dominant events at successively higher levels of tree structure.

The structural hierarchy of music rests on the perceived stability of the basic elements (either tones or chords). Those tones that are more stable are considered to be the structural elements of the piece such as the tonic and those closely related to it while the other, less related elements work to form the overall sense of expectation which is central to the musical experience (Patel, 2008). Figure 3 shows an example of structural hierarchy of a musical phrase.

Language, like music, is a syntactic system in that it has discrete elements (e.g. words) that can be combined through governing principles into hierarchical sequences (e.g. sentences: Patel, 2007). Figure 4 shows an example of the hierarchical structure of a sentence. Grammar is a finite set of combinatorial principles which is used to arrange a finite list of structural elements (Jackendoff, 2002). In language, words take the form of grammatical categories such as nouns and verbs, and these can take on grammatical functions such as subject, direct object, and indirect object within the structure of the sentence (Patel, 2003). Most native speakers are sensitive to and aware of the rules which govern the combination of these elements into structural form (Akmajian et al., 1984) as well as to violations of this grammar (Patel, 2003).

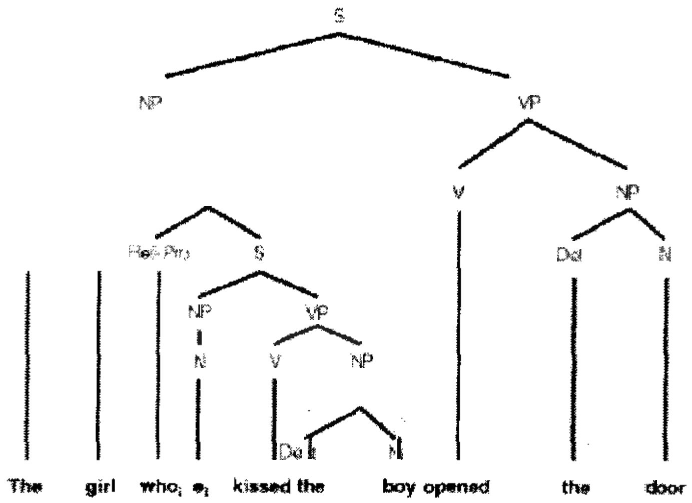

Figure 4 (Patel, 2007). A sentence of English, showing the hierarchical phrase structure

Mental representations of language and music syntax are quite different in that details of each system cannot be directly compared. Language has grammatical categories and functions which have no analog in music (Patel, 2007). Their similarities are found in the combinatorial rules that govern the hierarchical structure evident in both music and language (Patel, 2008).

Music Grammar: Experimental Evidence

Experienced perceivers, those who live in and are exposed to a culture's language and music, have implicit knowledge of the principles governing the grammatical structure of these two systems. They are able to detect violations within sequences, such as errors of agreement in language and sour notes in music. Most importantly, they are able to do so without training and within novel sequences (Patel, 2003).

Ayotte and colleagues (Ayotte, Peretz, & Hyde, 2002) used an anomalous pitch detection task in order to test participants' ability to detect "sour" notes within a melodic sequence. The task contained both familiar and novel melodies. Familiar melodies were well-known, cultural melodies such as folk songs while novel melodies were created by the experimenters. Half of the melodies presented contained a wrong note. The wrong note was created by modifying a note in the piece so that it no longer fit the harmonic context of the melody. Participants were asked to judge whether the melody contained a wrong note or not. Ayotte et al. found that participants were able to detect wrong notes at a level over 80%. Therefore, it was concluded that participants were sensitive to changes in the harmonic context.

Janata and colleagues (Janata, Birk, Tillman, & Bharucha, 2003) attempted to test the same ability in listeners but using, what they considered to be, a more natural musical context. That is, the stimuli were not short melodic trials which are not representative of normal music listening but rather consisted of a continuously modulating melody that lasted approximately 8 minutes. Within the melody were wrong notes that did not fit into the harmonic context. While the participants listened to the melody they were asked to press a key every time they heard a note that sounded "out-of place." Participants performed significantly above chance and were able to successfully respond when tones were unexpected within the harmonic context. The ability of the participants in these studies to detect notes that are unexpected within a harmonic context suggests that listeners have some implicit knowledge of the combinatorial rules of musical syntax and are sensitive to when this grammar is violated.

Language research uses two main methods to test participants' sensitivity to violations of syntax rules: self-paced reading tasks and event-related potential (ERP) recordings (e.g. self-paced reading: Chen, Gibson, and Wolf, 2005; McKoon &Ratcliff, 2007; Sieve et al., 2009; Vincenzi et al., 2003; ERP: Atchley et al., 2006; Eckstein & Friederici, 2006; Yamada & Neville, 2007; Vincenzi et al., 2003). Self-paced reading tasks typically consist of a sentence presented on a computer screen with one word or phrase bring presented at a time. These tasks are considered self-paced because the participant has control over the presentation rate of the words by pressing a key to advance to the next section of the sentence. The amount of time spent viewing each section is recorded and then a mean reading time for each section is calculated in order to look at the overall effect. In studies using self-paced reading time tasks to observe the effect of syntactic violations, a word is changed in order to create a syntactic anomaly. This word is known as the critical or target word of the sentence. Studies using this paradigm find that violations of grammar rules lead to increased reading times at the critical word (Ditman, Holcomb, & Kuperberg, 2007).

ERPs are on-line, continuous measures of brain electrical activity that occurs during a task, in this case language comprehension (Vincenzi et al., 2003). That is, while participants are reading sentences, ERPs are being derived from ongoing eletroencephalography (EEG) which are then correlated with behavioral data (Ditman et al., 2007). There are two main ERP components that have been indentified in relation to syntactic processing: an early left anterior negativity (ELAN) and a late centro-parietal positivity (P600) (Vincenzi et al., 2003). The ELAN has been proposed to be associated with the syntactic structure building process during which the structure of the sentence is being created based on word category information (Friederici & Meyer, 2004). The later P600 has been suggested to represent the processes associated with syntactic repair and reanalysis that occurs when a syntactic violation is present (Yamada & Neville, 2007). It has been suggested that the P600 may function in determining when the brain is processing a syntactic relation not predicted by the preceding structure (Patel et al., 1998). This P600 has been shown to be elicited by a number of grammar anomalies, including subject-verb agreement, case violations, and garden-path sentences (Vincenzi et al., 2003). Brain imaging studies have suggested that this P600 is generated by activity in the inferior frontolateral cortex which corresponds to the area of the brain known as Broca's area and therefore suggests that this area is involved in syntactic processing (Koelsch, 2006).

If music is also considered to have syntax which governs the combination of its basic elements, then violations of these rules should also elicit a P600. Patel and colleagues (1998) compared the ERP data of musicians during tasks involving language and music syntactic incongruities. The researchers found that the P600 elicited by language and music syntactic violations were statistically indistinguishable, which strongly suggests that the process that causes the P600 is not language specific. To further support this evidence, studies have elicited the P600 during music syntactic violations in non-musicians as well as musicians which therefore allows for greater generalizability of this phenomenon (e.g. Koelsch, Gunter, Friederici, &Schroger, 2000; Leino, Brattico, Tervaniemi, & Vuust, 2007).

Neuropsychology Evidence: Domain-Specificity

The general view in neuropsychology on the processing of music is one of domain-specificity; music is processed by a neural system that is separate from other auditory and language processing areas (e.g. Peretz & Coltheart, 2003). This domain-specificity is supported by neuropsychological case studies of individuals with selective deficits of music cognition after focal brain damage, which is known as acquired amusia, as well as individuals who suffer from congenital amusia, which is a life-long deficit of music processing. Patients who suffer from amusia perform at a normal level on speech and environmental sound tasks but not on tasks of music cognition (Poeppel, 2001). It is this functional dissociation between performance on music tasks and performance on speech and other auditory tasks that implies a specialization of neural networks for the processing of music.

Peretz and collegues (Peretz, 1993; Peretz et al., 1994) observed two patients in particular that exhibited this functional dissociation of music processing, C.N. and G.L.. C.N. suffered bilateral damage to the temporal lobe after two aneurysms, one of the right middle cerebral artery and one of the left middle cerebral artery. She complained of problems listening to music and that she could no longer recognized or remember familiar tunes. Much like C.N., G.L. suffered bilateral temporal damage due to strokes and also reported not being able to recognize familiar music as well as a loss of enjoyment in listening to music.

C.N, and G.L. participated along with a group of age, sex, musical background, and education matched controls in a series of experiments led by Peretz et al. (1994). Both C.N. and G.L. preformed at less than chance on most tasks of melody recognition and pitch discrimination and they were well below the range of the normal controls. Both patients were found to have severe deficits in music processing in relation to their performance on tasks concerning the processing of speech and environmental sounds. It was concluded that both C.N. and G.L. suffered from amusia without aphasia. That is, these individuals had severe deficits in music processing but their language production and comprehension abilities were unaffected. These findings support the notion that C.N. and G.L. sustained damage to the neural network which holds the domain-specific knowledge necessary for the processing of music and that music processing is neurally distinct from language processing (Peretz & Coltheart, 2003).

Research with congenital amusiacs also suggests this dissociation between music and language processing. Congenital amusia is a life-long deficit in music processing and it thought to affect approximately 4-5% of the population (Hyde & Peretz, 2004). Individuals who suffer from this disorder are unable to discriminate between pitches and cannot recognize familiar or popular songs along with many other music comprehension and production deficits. Ayotte and colleagues (Ayotte, Peretz, & Hyde, 2002) conducted a study with patients who suffer from congenital amusia in order to compare their functioning to normal controls on tasks involving not only music but language and environmental sounds as well.

On an anomalous pitch detection task, which tests patients' ability to detect a wrong note in both familiar and unfamiliar melodies, Ayotte et al. (2002) found that amusiacs performed close to chance and well below controls. On a recognition task involving familiar and unfamiliar melodies, lyrics, and environmental sounds those who suffered form congenital amusia performed at a level consistent with the normal controls for lyrics and environmental sounds. When asked to judge whether they recognized melodies however, amusiacs performed well below their matched controls and only slightly better than chance. In this manner, Ayotte et al. were able to demonstrate the selective musical deficits seen in congenital amusia as well as highlighting the dissociation between their performance on musical tasks and their performance on tasks involving language and environmental sounds.

While the cases of C.N. and G.L. and the congenital amusiacs involved musical processing deficits, other neuropsychological cases exhibit deficits of language processing at the phonological level. Poeppel (2001) reports on the phenomenon of pure word deafness (PWD). Those who suffer from PWD can no longer comprehend spoken material but they are able to process other sounds, including music. This deficit provides further evidence for a music-language dissociation.

Challenging Domain-Specificity

Although the neuropsychological evidence convincingly supports the theory of music and language as being independent and separate cognitive functions that are processed by domain-specific neural networks, neuroimaging research has challenged this view. Neuroimaging has shown a significant overlap in certain areas of music and language processing, especially syntactic processing, in normal individuals. The combined findings of ERP, EEG, MEG, PET, and fMRI studies suggest that music syntactic information is processed in areas once throught to be domain-specific to the processing of language (e.g. Brown, Martinez, & Parsons, 2006; Koelsch, 2006; Koelsch et al., 2002; Maess, Koelsch, Gunter, & Friederici, 2001; Patel et al., 1998).

This overlap of music and language processing was first demonstrated in an ERP study carried out by Patel et al. (1998). The researchers created both linguistic and musical stimuli which contained grammatical incongruities in the attempt to elicit the P600 in both music and language tasks. The P600, as discussed above, is normally related only to language syntactic processing. The musical stimuli consisted of musical phrases, some of which were manipulated to contain a target chord that was either from within the key of the phrase or from an outside key. As discussed previously, a chord from within the phrase of the key will be expected and therefore sound correct, while a chord from an outside key will be unexpected and will sound "alien" due to violating the grammar put forth by the earlier phrase structure (See Figures 1 and 2). The researchers found that the out of key chord elicited a P600 and that this P600 was statistically indistinguishable from the P600 which is generated by linguistic syntactic incongruities in sentences. These results suggest that there is some overlap in the processing of syntax in language and music.

Maess, Koelsch, Gunter, and Friederici (2001) used magnetoencephalography (MEG) in order to localize the neural areas associated with the processing of syntactic incongruities of music. The researchers presented participants with chord sequences, some of which contained harmonically unexpected chords. These harmonically unexpected chords represent a violation of music syntactic structure. Maess et al. found bilateral activation of the inferior part of Broadmann's area (BA) 44 when listeners were exposed to unexpected chords. Within the left hemisphere, this area is known as Broca's area. It has been proposed that language syntactic information is processed fast and automatically in Broca's area and its right homologue and, from the results reported by Maess et al., music syntax seems to be processed in a similar fashion in these brain areas.

An fMRI study carried out by Koelsch et al. (2002) investigated the brain areas activated during music processing. As in the studies described above, participants were presented with chord sequences, some of which contained unexpected musical events that violated musical syntax. These events activated a number of brain areas considered to be part of the neural network for language processing, including Broca and Wernicke's areas, the superior temporal sulcus, Heschl's gyrus, the anterior superior insular cortices, and both the planum polare and planum temporale. These findings support an overlap of music and language syntactic processing and further suggest that these areas are not as domain-specific as researchers had concluded based upon previous neuropsychology studies.

Patel and colleagues (Patel, 2005; Patel & Iverson, 2008) further support this overlap in processing through a study observing the musical syntactic processing abilities in agrammatic aphasia. Patel suggests that individuals who suffer from amusia-without-aphasia or aphasia-without-amusia have suffered damage to the areas of the brain that are related to the storage of domain-specific knowledge and representations but that the processing center remains intact (Patel, 1998). However, if this processing center is compromised, it is thought that these individuals will exhibit parallel deficits in linguistic and musical syntactic processing (Patel & Iverson, 2008).

Prior to the experimental tasks, participants were given a language pretest to assess their syntactic comprehension deficit as well as a music pretest in order to rule out any basic pitch perception or memory problems. The language pretest consisted of a sentence-picture matching task, of increasing syntactic complexity, in which participants were asked to listen to a sentence and then point to the corresponding picture. It was found that aphasic participants performed significantly worse than controls, with the difference increasing as the syntactic complexity of the sentences increased. Therefore, it was concluded that the aphasic individuals had a language syntactic comprehension deficit (Patel, 2005). The music pretest consisted of subtest from the Montreal Battery of Evaluation of Amusia. This item tests basic pitch perception and memory through a forced-choice, same-different task in which participants had to say that a pair of melodies were either the same or different from one another. The two groups did not differ in their basic pitch perception and memory performance.

Two experimental tasks were used to assess musical syntactic abilities of the aphasic participants. The first was both a linguistic and musical syntactic processing task. Participants were asked to listen to the presented sentences or melodies and to state whether the sequence was correct or if it contained an error. Half of the sentences contained either a semantic violation or a syntactic violation of subject-verb number agreement and half of the musical stimuli contained an out-of-key, harmonically unexpected chord. Aphasic participants performed worse than controls on both the linguistic syntactic task and in detecting harmonic anomalies in the musical sequences.

Also, performance on the musical syntax task was found to be a significant predictor of both aphasic and control participants' performance on the linguistic syntactic task. On the other hand, aphasic performed only marginally below controls on the linguistic semantic task and performance on the musical syntax task did not predict participants' performance on the linguistic semantic task.

The second task used was a harmonic priming task. This task tests music syntactic comprehension by observing the influence of a preceding harmonic context (e.g. the prime) on the processing of a target chord. That is, participants are first presented with a chord sequence and are then presented with a single chord (e.g. the target chord) and asked to judge if target is tuned or mistuned. Previous research suggests that the target chord is more easily and rapidly processed if it is close to the key implied by the prime. This ability reflects the implicit knowledge of the harmonic rules of musical syntax. The harmonic priming task showed a significant difference between aphasics and controls, in that controls showed normal harmonic priming while the aphasics failed to show a priming effect. This lack of a priming effect in aphasics suggests that they cannot activate the implicit knowledge of harmonic rules.

Patel and colleagues concluded that aphasics who have linguistic syntax comprehension problems may also have musical syntactic deficits. These results lend further evidence to the idea of a common process between language and music syntax comprehension and possibly a shared brain area that functions in syntax processing for both domains.

Theoretical Motivations: SSIRH and Statistical Learning

These results, along with neuroimaging findings, led Patel (2007) to create and define a model of resource sharing between language and music. He proposed that language and music share the neural resource that allows for the activation of networks that store domain-specific knowledge representations. This has been termed the "shared syntactic integration resource hypothesis" (SSIRH: Patel, 2007). This hypothesis predicts that simultaneous activation of the syntactic processing areas for both language and music stimuli should interfere with one another. That is, attention resources should be stretched between the two domains. This prediction has been supported through both neural and behavioral research (Koelsch, Gunter, Wittfoth, & Sammler, 2005; Sleve et al., 2009) Figure 5 shows a visual representation of SSIRH.

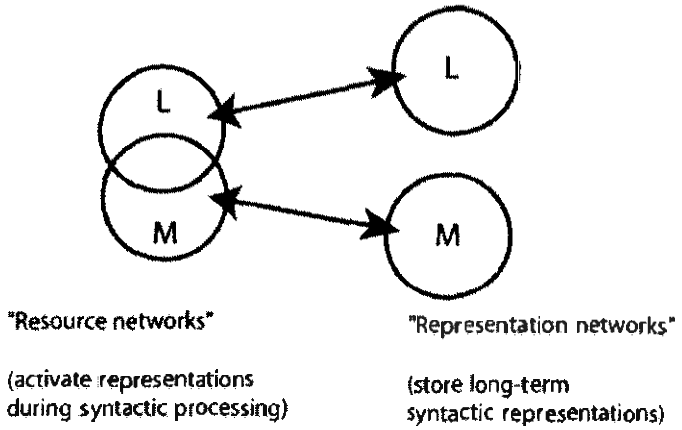

Figure 5 (Patel, 2007). Schematic diagram of the functional relationship between linguistic and musical syntactic processing. L = language, M = music.

Koelsch et al. (2005) observed an interaction between the processing of language syntax and the processing of music syntax using ERP data. They focused on the left anterior negativity (LAN) elicited by syntactic violations and the N400 elicited by semantically unexpected words. If language and music processing occur separately then the LAN and N400 should be unaffected by the presence of music-syntactic irregularities. Three types of sentence manipulations were used: syntactically expected/ semantically expected; syntactically expected/ semantically unexpected; and syntactically unexpected/ semantically expected. The target word was always the last word of each sentence. The musical stimuli consisted of musical sequences, half of which contained a music syntactically irregular or unexpected final chord. The sentences and music stimuli were paired together so that half of each sentence manipulation type was paired with an irregular musical sequence and that errors in both would simultaneously occur.

Koelsch et al. found that there is an interaction between language and music syntactic processing. When a linguistic-syntactically unexpected word was presented simultaneously with a music-syntactically irregular chord there was a reduction in the LAN. This interaction was not seen between semantic processing and music-syntactic processing. These results strongly support the assumption of a neural overlap between language and music syntax processing and therefore the SSIRH.

Sleve et al. (2009) used a behavioral measure, in the form of a self-paced reading task, in order to investigate the interaction between music and language processing. As previously described, self-paced reading tasks typically consist of small segments of a sentence being presented on the screen one at a time. Participants must press a key in order to move to the next segment and the amount of time spent looking at each segment is recorded. Sentences often contain a critical point, where a syntactic, semantic, or other language error may appear, and it is at this point where differences between manipulations were expected to be.

Slevc and colleagues (2009) used both semantic and syntactic violations within their language stimuli. Semantic violations were created by including a word that was inconsistent with the sentence context. For example: *The boss warned the mailman to watch for angry **pigs** when delivering the mail.* The semantically inconsistent word (*pigs*) leads to a semantic violation due to the unexpected nature of the word based on the preceding sentence context. In reading tasks, semantic violations generally appear as a disruption of or an increase in reading time (Vincenzi, et al, 2003).

Syntactic errors were created through the use of garden-path sentences. These sentences are considered temporarily ambiguous, in that readers adopt an analysis only to later find that analysis to be ungrammatical, which in turn leads to processing difficulty (van Gompel, Pickering, Pearson, & Jacob, 2006). An example of a garden-path sentence used by Slevc et al. (2009) is: *The scientist wearing thick glasses confirmed the hypothesis **was** being studied in his lab.* In these garden-path sentences the word *that* is omitted and this omission causes a violation of syntactic expectancy. The sentence including the word *that* is much simpler to read: *The scientist wearing thick glasses confirmed <u>that</u> the hypothesis **was** being studied in his lab.* Previous studies have shown that the processing of this type of sentence is more difficult as evidenced in longer reading times for *that*-less sentences in comparison to those which contain *that* (Trueswell, Tanenhaus, & Kello, 1993; van Gompel et al., 2006).

Music stimuli consisted of simple Western tonal chord sequences. Half of these melodies contained an unexpected chord that did not fit the harmonic context and therefore violated musical grammar. These melodies were then paired with both language manipulations so that the critical points of the sentence were paired with the critical point (where the unexpected chord would occur) of the musical sequences. Figure 6 is a visual representation of the experimental procedure used by Sievc et al.

Sievc et al. found a significant three-way interaction between linguistic manipulation (syntactic or semantic), language expectancy (expected or unexpected), and music expectancy (expected or unexpected). Those words that were either syntactically or semantically unexpected were read more slowly that their expected counterparts. Hearing a simultaneous unexpected chord caused even greater slowing when combined with syntactically unexpected words, but not when paired with semantically unexpected words. Therefore, the processing of music-syntactic errors interfered with the processing of syntactic, but not semantic, elements of language. These results further support the SSIRH in that is seems that neural resources were being shared between language and music syntactic processing.

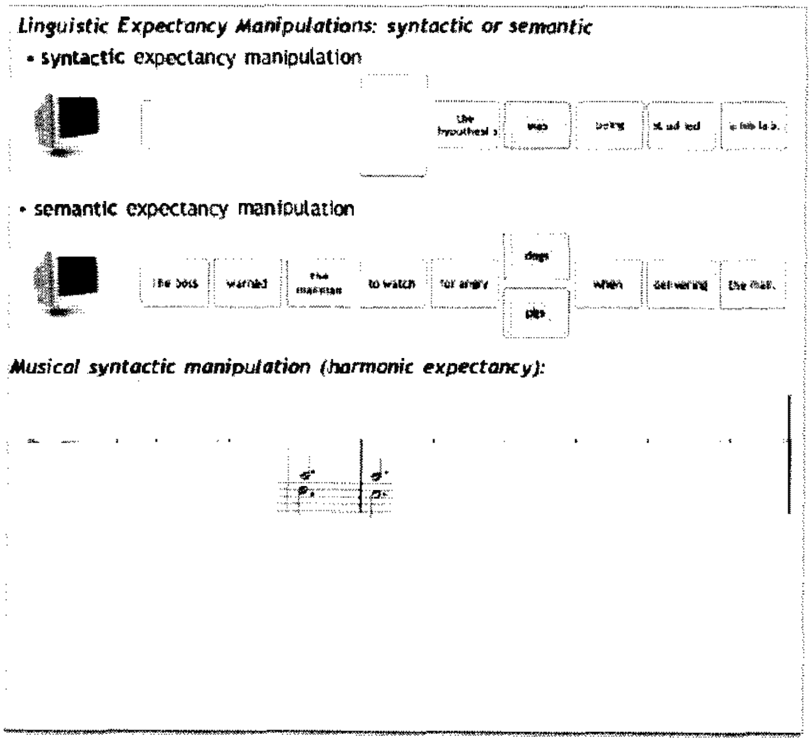

Figure 6 (Slevc et al., 2009). The top portion of the figure represents the procedure for displaying the segments of sentences on the screen during the self-paced reading task. The bottom shows an example of a chord sequence that would be paired with the sentences.

In the SSIRH, violations of syntactic expectancy in both language and music are used in order to observe interference between the processing of the two. Syntax can be defined as the principles governing the combination of discrete structural elements (Jackendoff, 2002). In music the structural elements are tones and in language, words (Patel, 2003). These combinatorial principles define the way in which these elements are combined into larger structures: phrases and sentences in language and keys and melodies in music (Patel, 2003). These structures, both in music and in language, are generated hierarchically starting from the basic units (Brown, Martinez, & Parsons, 2006).

Research discussed earlier supports the idea that an individual does not have to be a trained musician in order to have knowledge of the rules that govern musical syntax, just as one does not need to be a linguist in order to have knowledge of the rules that govern language syntax (Koelsch et al., 2000). The implicit knowledge of both language and music syntax is acquired early in development without explicit tutoring but through exposure and experience with a specific culture's rules (Peretz, 1993).

The ability of humans to acquire this knowledge implicitly has been studied using a statistical learning approach in which participants are exposed to a novel grammar (e.g. Creel, Newport, & Aslin, 2004; Loui & Wesssel, 2006; Saffran, 2001, 2002; Saffran et al., 1999, 2008; Thompson & Newport, 2007). Statistical learning is the process of detecting patterns and using these patterns to discover the underlying structure. Both music and language contain a number of cross-cultural statistical regularities including the use of a hierarchical phrase structure (Saffran, 2003). The research with statistical learning focuses on human's ability to pick up on these structural regularities and use them to create an implicit knowledge of the governing principles of the grammar which they are exposed to (Saffran, 2002).

Saffran (2001, 2002) has tested both children and adults in their ability to acquire the basic knowledge of an artificial language using a statistical learning approach. Saffran used an artificial grammar which contained predictive dependencies; that is, the presence of some word categories relies on the presence of others (Saffran, 2002). This is similar to the underlying structure of natural languages. Saffran suggests that through these predictive dependencies, knowledge of hierarchical phrase structure is formed. Adults and children were exposed to this type of artificial grammar and then during the test phase were asked to judge pairs of sentences on the basis of which sounded more like the exposed grammar (Saffran, 2001, 2002). Overall, Saffran found that both adults and children are able to learn the grammar rules of the artificial languages that contained predictive dependencies.

Music, like language, contains statistical regularities that form a hierarchical phrase structure and therefore music syntax may also be acquired through the use of statistical learning. Saffran and colleagues (McMullen & Saffran, 2004; Saffran, 2003; Saffran et al., 1999) have considered the possibility of an overlap in mechanisms for the learning of syntactical knowledge in both language and music due to the similarity of their culture specific rules and their combinatorial nature. Saffran et al. (1999) developed a tone-language based on an earlier artificial language grammar, in that tones replaced each syllable of the language grammar. Participants were exposed to the tone-language and were then given a forced-choice task during which they had to select the tone sequence that sounded the most familiar. Participants were able to correctly identify the tone sequences from the artificial tone-language grammar. This suggests that the same statistical learning mechanism may be in use for both language and music. However, there were a few limitations to this study. One is that they used Western scale tones which may interfere with learning due to the fact that participants most likely have some implicit knowledge defining the organization of these tones. Another possible limitation is the lack of musicality of the tone sequences (Saffran, 2003).

In order to extend the statistical learning approach to music acquisition, an artificial grammar must be created without the use of the Western music idiom. A possible means through which to develop an artificial musical grammar is proposed in a study by Loui and Wessel (2006). The researchers defined two new musical grammars using the Bohlen-Pierce (BP) scale. The BP scale is a microtonal tuning system which contains 13 tones and creates a tritave scale (Figure 7). This scale differs from the Western tonal scale system in a number of ways. The most obvious is that the Western system contains 12 tones while the BP system contains 13. The other difference is that the Western system has a 2:1 frequency ratio while the BP system has a 3:1 frequency ratio. This ratio difference is what causes the BP scale system to sound so distinct and completely unlike the Western tonal system.

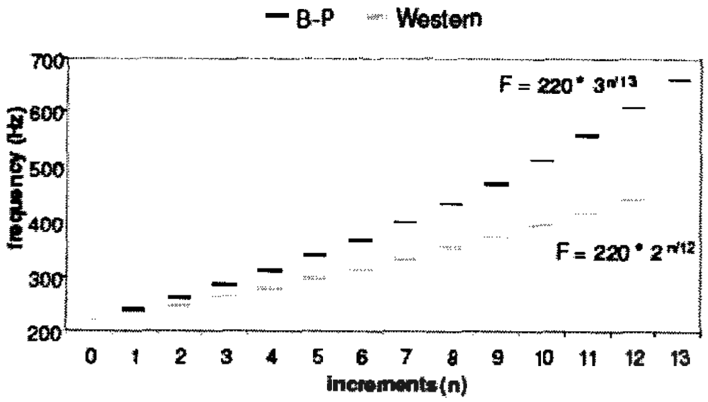

Figure 7 (Loui & Wessel, 2006). Frequencies along the Bohlen-Pierce scale and the Western scale.

Using this scale, Loui and Wessel (2006) were able to compose two different chord progressions which each consisted of four chords with three notes per chord (Figure 8). Each chord progression defined a grammar, each with its own set of grammatical rules based on predictive dependencies. Sets of melodies were composed using the chord progressions as a basis (Figure 9). Each melody ranged from four to eight notes and was unique to its grammar.

	Pitch number			
Grammar I	10	7	10	10
	6	4	7	6
	0	0	3	0
Grammar II	10	10	7	10
	6	7	4	6
	0	3	0	0

Figure 8 (Loui & Wessel, 2006). These chord progressions formed the bases of the two sets of grammatical rules which were used to form the stimuli in Loui and Wessel (2006). By applying these rules in the fashion of a finite-state grammar, they composed sets of melodies from the above chord progressions.

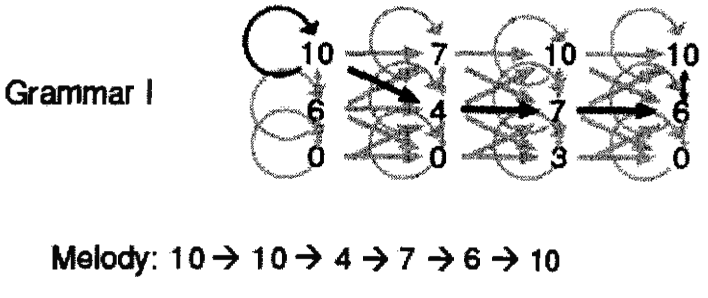

Figure 9 (Loui & Wessel, 2006). An illustration of applying a finite-state grammar as a set of rules to compose one melody based on finite-state Grammar I. Dark arrows illustrate the paths taken, whereas light arrows illustrate the other possible paths that are legal in the grammar. The resultant melody is shown at the bottom of the figure.

Loui and Wessel (2006) ran three experiments in order to observe the extent to which participants would learn the presented grammar with the third experiment being the most pertinent to this research. In this experiment, 24 undergraduate students with at least five years of musical experience were randomly assigned to one of the two grammars. Five-hundred melodies, each eight notes in length, were composed using the two grammars. The participants were first asked to complete a pre-exposure probe-tone task in which they heard a melody followed by a single tone. There were asked to rate how well that tone fit the presented melody. Next, participants were exposed to 400 melodies which were presented at random with no repeats for 30 minutes while participants completed a drawing task used to alleviate boredom.

After exposure, participants were given a forced-choice recognition and generalization task. During this task participants heard two melodies and were asked to select which melody sounded more familiar. The recognition task consisted of two melodics, one from the exposed grammar and one from the unexposed grammar. The generalization task used a novel melodies composed using the exposed grammar paired with a melody from the unexposed grammar. A post-exposure probe-tone task was then given, which was identical to the pre-exposure task.

Overall, Loui and Wessel (2006) found that when participants were exposed to a large number (400) of melodies that they were able to learn that statistical regularities of the grammar. Participants performed significantly above chance on the recognition and generalization forced-choice task. That is, participants were not only able to recognize previously heard melodies but they were also able to use their knowledge of the grammar to generalize to other novel melodies from that grammar. Participants also performed significantly better on the probe-tone task after being exposed to the grammar than before exposure. The results suggest that participants were able to learn the grammar and were sensitive to the underlying statistics of the grammar. This research supports the use of this artificial musical grammar as a means through which to study the implicit learning of new music-grammatical structures.

These results relate back to the idea proposed by McMullen and Saffran (2004) of a possible shared mechanism for the learning of both music and language. The parallel findings of Loui and Wessel (2006) and linguistic statistical learning studies (Saffran 2001, 2002) show that humans are able to use statistical regularities to learn the rules of artificial grammars. These shared learning mechanisms further support the hypothesis of an overlap in the syntactical processing of music and language.

The present study attempted to replicate and extend the findings of both Loui & Wessel (2006) and Sieve et al. (2009). Participants were exposed to an artificial music grammar similar to that used by Loui & Wessel (2006) in order to investigate the ability of non-musicians to acquire the new grammar in comparison to musicians. Participants were expected to learn the new grammar and this learning was not expected be dependent on their prior musical knowledge; that is, no difference in performance was expected between musicians and non-musicians on tasks that test their learning of the artificial music grammar.

In this study, the artificial grammar was implemented into the SSIRH procedure used by Slevc et al. (2009) along with the original Western grammar and a random control grammar. An interaction between language syntactic-unexpectancy and music syntactic-unexpectancy was expected for the Western condition such that an increase in reading time when these violations was predicted when these violations were paired. This interaction was also expected in the Acquired condition if participants were able to successfully learn the artificial grammar. Such results would provide further evidence for the SSIRH and for the use of the artificial grammar in studying music acquisition.

Overall, there were two main research goals for the present study: 1) Observe non-musicians' ability to learn the artificial music grammar; 2) By using a SSIRH paradigm, further test the learning of the new music grammar as well as explore the possibility that shared resources between music and language extend to the processing of novel music grammars.

Methods

Participants

One hundred fifty-one Seton Hall University undergraduates (95 female) participated in exchange for course credit. Close to ninety percent (89.4%) reported receiving less than 5 years of musical training, with 61.6% receiving a year or less. The other 10.6% received 6+ years of training and only one participant identified themselves as majoring in music. Therefore, the majority of participants were classified as non-musicians. Each of the participants was randomly assigned to one of the three musical grammar conditions.

Materials

Twenty-four critical sentences, in which either the syntactic or semantic expectancy were manipulated, were adapted from Sleve et al (2009). Twelve sentences contained syntactic expectancy manipulations which were achieved through the use of garden path sentence structures. Sentences with syntactic errors contained a reduced sentence complement (the word *that* is omitted) while those without the errors included a full sentence complement (contains the word *that).* The syntactic interpretation of the sentences was therefore either unexpected or expected, respectively, at the critical word (underlined below). The other 12 sentences contained semantic expectancy manipulations. These manipulations were achieved through the use of either a semantically inconsistent or consistent word (underlined below). Depending on the consistency of the word used, the semantic interpretation was either unexpected or expected at the critical word (underlined below). An additional 24 sentences were used as filler sentences and contained neither syntactic nor semantic errors. These filler sentences were not used in the analyses.

Syntactic expectancy manipulation example:

The scientist confirmed (that) the hypothesis <u>was</u> being studied in his lab.

Semantic expectancy manipulation example:

The farmer went to the barn to milk the large (<u>cows/cats</u>) before dawn.

Filler sentence example:

When the monster suddenly appeared, the audience shrieked.

A separate musical sequence was paired with each sentence. Musical stimuli consisted of three musical grammars: Western, Acquired, and Random. The Western music grammar was created through the use of a compositional style known as counterpoint. This style creates simple, single note melodies. The melodies ranged in length from 5 notes to 11 notes in order to correspond to the sentence lengths. They were recorded using a piano timbre. Twenty-four of the melodic sequences were manipulated to contain a note from a distant key which created music syntax errors. These music syntax errors were occurred simultaneously with the critical point in the sentences.

The Acquired grammar consisted of 500 melodies created from the Grammar I defined by Loui & Wessel (2006). Like the Western melodies, the acquired melodies range in length from 5 notes to 11 notes and were recorded using a piano timbre. There were no constraints on interval size within the melodies which follows the methods presented by Loui and Wessel (2006). To create syntactic violations, an out of grammar tone that violates the dependent phrase structure was inserted at a point that matched with the critical word in the sentences. This insertion should be heard as being similar to inserting an out-of-key tone into a Western tonal sequence. Overall 48 melodies were manipulated, with 24 containing music-syntactic manipulations, to be used in the test phase.

The Random, non-grammar musical stimuli consisted of 500 randomly generated sequences created using the 13 tones of the Bohlen-Pierce tuning. These sequences ranged in length from 5 notes to 11 notes and were recorded using a piano timbre. Forty-eight of these sequences were used in conjunction with the sentences and 24 had a note changed at the critical point. Because these sequences had no grammar, a change was not considered to be a syntactic manipulation.

Procedure

Each participant was tested individually. They were seated in front of a computer and asked to wear a pair of headphones during the experiment. Participants began with a pre-exposure probe-tone task during which they heard a melody followed by a single tone. They were asked to rate how well that tone fit the melody using a 7 point scale, with 1 being the least fitting and 7 being the best fitting. Following the probe-tone task participants were exposed to one of the three grammars. Four hundred melodies were presented in random order with no repeats and were heard for approximately 30 minutes. While listening to these melodies, participants were given the option of sitting quietly or completely a coloring task in order to alleviate boredom.

Following exposure, participants were asked to complete a forced-choice recognition and generalization task. A total of 20 trials were given to each participant, 10 were the recognition task and the other 10 were the generalization task. The recognition task contained 10 melodies taken from the exposed grammar paired with 10 melodies taken from an unexposed grammar of the same type (i.e. Western exposed melodies were paired with Western unexposed melodies). The generalization task consisted of 10 novel melodies from the exposed grammar paired with 10 melodies from the unexposed grammar. The participants were asked to indicate which melody (1 or 2) sounded more familiar to them based on what they had heard during the exposure phase. This phase was followed by a post-exposure probe-tone task which was exactly the same in procedure to the pre-exposure probe-tone task.

Participants then completed a sentence-reading task. Short segments of each sentence appeared in the middle of the computer screen and participants were instructed to press a key in order to move on to the next segment. The tone sequence was played over headphones and each segment of the sentence was accompanied by a tone. The complete sentence formed a complete musical sequence. The tone began with the appearance of the segment and ended when the participant pressed the key to move on. If the participant did not hit the key the tone decayed over 1.5 seconds. Each complete sentence was followed by a comprehension question with a yes/no response (e.g. *Did the farmer get up early?*). Correct responses to the questions caused a screen with *Correct!* to appear while an incorrect response elicited *Incorrect!,* both of which remained on the screen for 1500ms followed by the start of the next sentence. The participants were asked to read the sentences as quickly as possible while retaining accuracy on the comprehension questions. They were told that they would hear a tone with the presentation of each sentence segment but to focus on the sentences not the melodies. Response times, time spent looking at each segment, were recorded. Each participant completed a total of 48 trials.

The session concluded with a survey containing some demographic questions as well as questions about their music experience (see Appendix).

Design and Analysis

The sentence-reading task experimental design consisted of one between-participant factor with 3 levels (music exposure type: Western, Acquired, and Random) and three within-participant factors with two levels each (language expectancy: expected or unexpected), language type: syntactic or semantic, and music expectancy: expected or unexpected). Each participant saw a given sentence only once. The conditions were counterbalanced so that the manipulation of a sentence was rotated; in that, for a given sentence each condition contained a different within-item manipulation of language and music expectancy of that sentence. The items were presented in a pseudo-random order provided by Sleve et al. (2009) in which critical and filler sentences were presented on alternate trials.

The sentence-reading data were cleaned following the procedure outlined by Sieve et al. (2009). Reading times (RTs) below 50ms or above 2500ms per segment were discarded along with RTs that were above or below 2.5 standard deviations from each participant's mean reaction time. These criteria led to exclusion of 1.3% of critical point observations. It is at the critical point where violations in both language and music could occur. RTs were logarithmically transformed. These reading times were analyzed using a mixed between-within subjects analysis of variance (ANOVA). Three sections of the sentences were analyzed: the critical point, the section immediately pre-ceding the critical point (pre-critical) and the section immediately following the critical point (post-critical). Separate analyses were performed for each of these three sections.

These analyses were performed on just the Western grammar in order to look for a possible replication of Slevc et al.'s findings. It was expected that errors in Western music-syntax would interact with language-syntax errors but not with language-semantic errors. Specifically, this interaction between simultaneous music- and language-syntax errors was predicted to cause an increase in reading times greater than those seen with errors in language-syntax alone.

These analyses were then performed again on the Acquired and Random grammars together in order to isolate the effects of the Acquired grammar on reading times. Similar results to those expected with the Western music grammar were anticipated for the Acquired group if participants were able to learn the grammar. No interaction was expected between language-errors and music-errors in the Random condition due to the lack of syntactical regularities within these sequences.

The music tasks were also analyzed in order to look for the rate of performance across grammars and also to determine possible learning of the Acquired grammar. The forced-choice task was analyzed using a one-way between-groups ANOVA for the recognition pairs and again for the generalization pairs. It was expected that participants in the Western music group would perform above chance on all tasks because of their life-time experience with the Western music-syntax. If participants were able to learn the artificial music grammar, it was expected that their performance on the forced-choice tasks would be similar to those seen in the Western music group. Participants in the random condition were expected to perform at chance on the forced-choice task because of the lack of syntactical regularities with in the tone sequences. Data from the probe-tone tasks were not analyzed due to observations that this task was too difficult for participants, in that, participants expressed confusion about the task and this confusion was displayed through the inconsistency of their responses.

The effects of gender were considered for both the ability to learn the newly acquired grammar and for performance on the sentence-reading task. Gender was considered based on previous research that suggests differences between males and females in both language and music processing. Females have been found to perform better on language performance tasks and also show better first-language acquisition early in life than their male counterparts (Burman, Bitan, & Booth, 2008; Wallentin, 2009). Evidence has also suggested that gender plays a role in second-language learning; females perform better than males on both syntactic and semantic tasks in the second-language (Andreou, Vlachos, & Andreou, 2005). There are also differences in brain activation for language processing between males and females. Neuroimaging studies have observed left lateralized activation in males during language tasks while females show a more bilateral pattern of activation (Shaywitz et al., 1995).

Few studies have examined at gender differences in music processing, despite its similarity to language processing. However, a recent brain imaging study by Koelsch and colleagues (Koelsch, Maess, Grossman, & Friederici, 2003) found brain activation differences during music processing tasks. It was observed that males showed a clear right hemispheric pattern of activation while females showed a bilateral pattern of activation.

Based on this previous research, it was anticipated that music processing may be influenced by similar sex differences. Therefore, gender was expected to affect both the ability to learn the new grammar and performance on the sentence reading task. Due to observations that females are better at learning a second language, it was expected that females would perform above their male counterparts on the tasks which test their learning of the new music grammar. Gender was predicted to possibly affect the sentence reading task in two ways: 1) If females were in fact able to better learn the new grammar, they would be more sensitive to errors within the new music grammar and therefore would show greater reading times when simultaneous music- and language-syntax violations occurred; 2) Since the SSIRH paradigm explored the possible overlap of syntax processing mechanisms between music and language, it was possible that results would be affected by the differences in brain areas activated during music and language processing observed between the genders. It would follow, then, that if females processed both music and language bilaterally, there would be a greater chance of processing mechanisms overlapping and therefore a greater chance of an interaction between music- and language-syntax processing for women in the study.

Also, the possible relationship between ability to learn the Acquired grammar and the performance on the reading task was examined. Participants in the Acquired group that performed well above chance on both of the forced-choice task are assumed to have more completely learned the syntactical regularities of the new music grammar. Because of their more precise representation of these regularities, it was possible that these participants would possibly be more sensitive to violations of the new music grammar. Therefore violations of the new music grammar would lead to greater interaction with language-syntax violations in the sentence-reading task.

Results

Figures 10 and 11 show results from the recognition and generalization tasks, respectively. It was predicted that participants in the Western condition should perform above chance on these tasks due to their life-long exposure to the Western music grammar. Also expected were similar results for the Acquired group if they had successfully learned the regularities of the new music grammar during the exposure phase. Those in the Random group were expected to perform below chance because of the lack of syntactical regularities within the tone sequences.

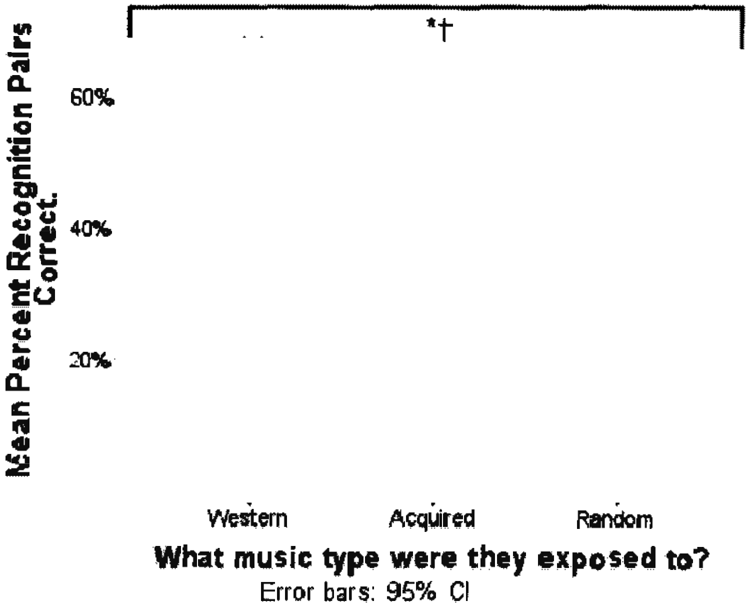

Figure 10. The Number of Recognition pairs correct, with the chance level marked at 50%.
Note: * significantly above chance at the p<.05 level and † significantly different from the Random condition at the p<.05 level.

Forced-choice recognition tests demonstrated that participants in the Western and Acquired conditions were able to successfully recognize previously presented melodies. Both of these groups performed significantly above chance: Western, $t(51)=3.25, p<.01$ and Acquired, $t(51)=5.5, p<.001$. However, those in the Random condition performed at chance on this task. There was a statistically significant difference between the three music grammar groups on this task (*F*(2, 148)=8.7,*p*=.00). Post-hoc comparisons using the Tukey HSD test indicated that there was no difference between the Western (*M*=5.8, *SD*=1.8) and Acquired (*M*=6.2, *SD*=1.6) conditions mean scores but those two conditions' mean scores were significantly different from the Random condition's (*M*=4.8, *SD*=1.6).

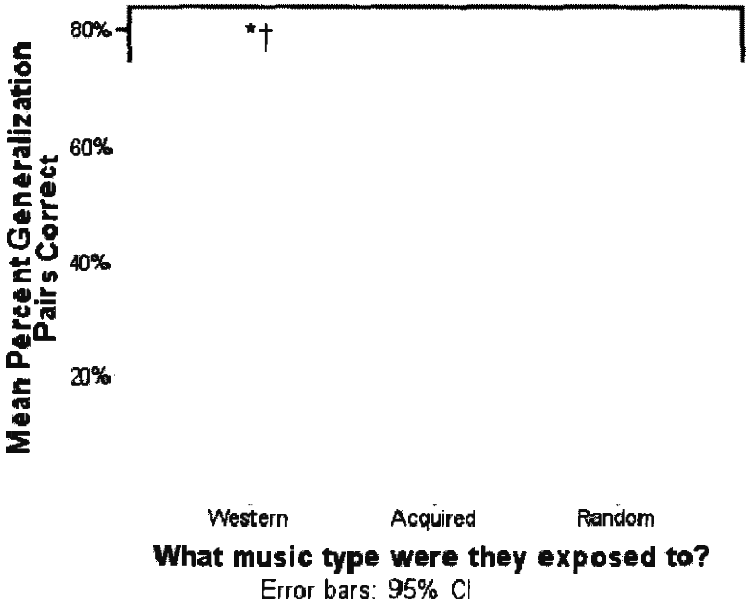

Figure 11. The Number of Generalization pairs correct, with the chance level marked at 50%. Note: * significantly above chance at the p<.05 level and † significantly different from the Acquired and Random conditions at the p<.05 level.

The results for the forced-choice generalization tests showed a different pattern of results in that only those in the Western condition performed significantly above chance on this task: *t*(51)=6.92,*p*<.001. Again, there was a statistically significant difference between the three music grammar groups (*F*(2, 148)=19.8, *p*=.00). Post-hoc comparisons using the Tukey HSD test demonstrated that there was no difference between the mean scores of the Acquired (*M*=5.1, *SD*=1.5) and Random (*M*=5.0, *SD*=1.4) conditions. However, mean scores for both the Acquired and Random conditions were significantly different from the mean scores of the Western condition (*M*=6.8, *SD*=1.9). These forced-choice tests showed that participants in the Acquired condition were able to recognize those melodies presented during exposure but were unable to generalize their familiarity to novel melodies from the exposed grammar.

Gender was found to have no affect on participants' performance on the forced-choice tasks for the Acquired music grammar (Recognition: *F*(2, 51)=.24, *p*=.78; Generalization: *F*(2, 51)=2.16, *p*=.13). It had been expected that, due to their superior performance on second-language learning tests, females would better learn the new music grammar and therefore perform above than their male counterpoints on the music tasks. However, the results found may be based on the skewed number of females (34) to males (17) in the Acquired group.

The RTs from the sentence reading task were analyzed using separate ANOVAs for the three sentence regions being considered: pre-critical, critical, and post-critical. These analyses were performed for the Western grammar condition and again for the Acquired and Random conditions together. Increases in reading time at the critical points when simultaneous music- and language-syntax violations occurred were expected for the Western and Acquired conditions. These results were expected for the Western condition because participant should have a complete syntactical knowledge of the regularities of this musical system because of life-long exposure and should therefore be sensitive to violations of the Western music grammar. Similar results should be seen with the Acquired group if they were able to learn and form a representation of the syntactical regularities of the new music grammar. No increase was expected at the critical point when simultaneous music- and language-syntax errors occurred for the Random condition due to the lack of syntactical regularities and therefore no learning of the syntax and no Random music-syntax errors within the music sequences could occur

Table 1. Mean reaction times (in milliseconds, with standard errors in parenthesis) by sentence region and by condition. Pre-Critical region is defined as the sentence segment directly preceding the critical point and the post-critical region is the sentence segment directly following the critical region. The critical region of the sentence is where the language or music manipulations could occur.

Pre-Critical Region						
	Syntactically			**Semantically**		
	expected	unexpected	difference	expected	unexpected	difference
Western Expected	639 (33)	603 (29)	-36	478 (17)	479 (23)	1
Western Unexpected	626 (32)	555 (28)	-71	483 (22)	478 (21)	-5
Acquired Expected	617 (39)	622 (39)	5	516(31)	506 (28)	-10
Acquired Unexpected	646 (41)	604 (38)	-42	505 (30)	501 (25)	-4
Random Expected	604 (41)	588 (41)	-16	488 (32)	492 (29)	4
Random Unexpected	606 (43)	571 (40)	-35	497 (32)	472 (26)	-25
Critical Region						
	Syntactically			**Semantically**		
	expected	unexpected	difference	expected	unexpected	difference
Western Expected	515 (22)	524 (26)	9	481 (20)	509 (26)	28
Western Unexpected	498 (18)	519 (27)	21	488 (26)	515 (26)	27
Acquired Expected	525 (33)	577 (40)	52	515 (31)	556 (40)	41
Acquired Unexpected	535 (36)	570 (35)	35	526 (32)	556 (38)	30
Random Expected	506 (35)	530 (42)	24	452 (33)	524 (42)	72
Random Unexpected	513 (38)	524 (37)	11	468 (34)	478 (39)	10
Post-Critical Region						
	Syntactically			**Semantically**		
	expected	unexpected	difference	expected	unexpected	difference
Western Expected	512 (24)	527 (24)	15	506 (20)	583 (30)	77
Western Unexpected	520 (23)	541 (31)	21	515(25)	565 (21)	50
Acquired Expected	527 (33)	576 (41)	49	533 (30)	606 (36)	73
Acquired Unexpected	536 (37)	566 (32)	30	531 (29)	594 (35)	63
Random Expected	491 (35)	523 (43)	32	477(31)	571 (38)	94
Random Unexpected	503 (39)	512 (34)	9	486(30)	555 (37)	69

The analysis for the Western grammar condition consisted of a 2 (language type) X 2 (language expectancy) X 2 (music expectancy) within-participant ANOVA. The Acquired and Random analysis was the same except for the addition of a between-subject factor (music type). Table 1 shows the mean RTs for each condition by sentence region. For the sake of brevity, only significant and marginally significant results of the ANOVAs are reported. The results from the Western condition will be considered first.

In the pre-critical region, RTs were longer for the syntactically-manipulated sentences than for the semantically-manipulated sentences (a main effect of language type: $F(1, 51)=214.42, p<.001, \eta^2_p=.81$). Also, RTs were longer for linguistically expected conditions than for unexpected conditions (a main effect of language expectancy: $F(1, 51)=8.90, p<.01, \eta^2_p=.15$).Slevc et al. (2009) also found this main effect of language expectancy. Slevc et al. explained this difference in RTs between the expected and unexpected conditions as being representative of earlier differences in the sentences (e.g., the presence or absence of *that*). Overall, the expected conditions for syntactically-manipulated sentences had the longest RTs (an interaction between language type and language expectancy: $F(1, 51)=3.98, p=.05, \eta^2_p=.07$). There was also a main effect of music expectancy ($F(1, 51)=6.19, p<.05, \eta^2_p=.11$) in that expected music conditions had longer RTs than unexpected music conditions.

In the critical region, RTs were longer for syntactically-manipulated sentences than for semantically manipulated sentences (a main effect of language type: $F(1, 51)=9.64, p<.05, \eta^2_p=.16$). No other effects reached significance.

In the post critical region, RTs were longer for semantically-manipulated conditions than for syntactically-manipulated (a main effect of language type: $F(1, 51)=7.57, p<.01, \eta^2_p=.13$). Also, RTs were slowed by both syntactic and semantic unexpectancy (a main effect of linguistic expectancy: $F(1, 51)=9.47, p<.01, \eta^2_p=.16$). This was especially true for semantically manipulated sentences (an interaction between language type and language expectancy: $F(1, 51)=5.70, p<.05, \eta^2_p=.10$). No other effects reached significance.

The pre-critical, critical, and post-critical regions of the Acquired and Random conditions were then considered. In the pre-critical regions, syntactically-manipulated sentences had longer RTs than the semantically-manipulated sentences (a main effect of language type: $F(1, 97)=249.67, p<.001, \eta^2_p=.72$). No other effects were found to be significant.

In the critical region, RTs were longer for syntactically-manipulated sentences than for semantically manipulated sentences (amain effect of language type: $F(1, 97)=28.8, p<.001, \eta^2_p=.23$). This was especially true in Acquired grammar conditions (an interaction of music type and language type: $F(1, 97)=4.92, p<.05, \eta^2_p=.05$). Also, RTs were longer for linguistically unexpected conditions (a main effect of language expectancy: $F(1, 97)=13.98, p<.001, \eta^2_p=.13$). There was also a trend towards a significant three-way interaction between music type, language type and language expectancy ($F(1, 97)=3.53, p=.06, \eta^2_p=.04$) in that unexpected syntactic manipulations in the Acquired grammar conditions had the longest RTs. No other effects reached significance.

In the post-critical region, RTs were longer for semantically-manipulated than for syntactically manipulated sentences (a main effect oflanguage type: $F(1, 97)=6.37, p<.05, \eta^2_p=.06$). Also, RTs were slowed by linguistic unexpectancy (a main effect of linguistic expectancy: $F(1, 97)=46.59, p<.001, \eta^2_p=.32$). Overall, the unexpected semantically-manipulated sentences had the longest RTs (an interaction oflanguage type and language expectancy: $F(1, 97)=9.14, p<.01, \eta^2_p=.09$).

Performance on the forced-choice recognition and generalization task, where participants were asked to select the most familiar melody out of two presented melodies, was found to have no effect on the sentence reading task for both the Western and Acquired conditions (F(1, 149)=.89,p=.35, $\eta^2{}_{,p}$=.017). It had been predicted that those who performed well above chance on these music tasks may have a more complete representation of the syntactical regularities of the music grammar and may be more sensitive to violations of the music grammar which would result in a greater increase in response time at the critical point of the sentence. However, these results could be due to the small number of participants with greater than 60% correct on those tasks (Western N = 22, Acquired N=18).

Gender was expected to affect performance on the sentence-reading task. It was predicted that females may show a greater interaction between music-syntax and language-syntax violations. This prediction was considered because of the bilateral brain activation seen during both music and language processing tasks in contrast to males more hemispheric-specific activation. Gender did not affect the findings for the sentence reading task when looking at the Acquired and Random conditions. However, gender did affect the findings at the critical point for the Western conditions.

There was a significant four-way interaction between gender, language type, language expectancy, and music expectancy (F(1, 50)=4.6,p<.05, η^2_p =.08). Females had overall longer RTs for all conditions. Most interesting was that females had longer RTs for syntactically unexpected conditions paired with an unexpected music-manipulation (M=575.48, SE=33.7) than for syntactically unexpected conditions paired with an expected music-manipulation (M=542.93, SE=33.75). This is consistent with findings of

Slevc at al. (2009) in that the longest RTs in the sentence-reading task occurred at the critical point during which simultaneous errors in music- and language-syntax occurred. Males on the other hand, showed the opposite direction of RTs (syntactically unexpected and music unexpected: *M*=441.47, *SE*=39.35 and syntactically unexpected and music expected: *M*=498.91, *SE*=39.42).

Discussion

The present study tested the ability of participants to learn an artificial music grammar as well as implementing this grammar into a SSIRH procedure in order to explore a possible interaction with language processing. During the first phase of the experiment, participants were exposed to a novel musical grammar and were given a series of tasks to test their ability to learn the new grammar. Participants were able to recognize melodies they had heard during the exposure phase but were unable to generalize their learning to novel melodies from that grammar.

The second phase of the experiment consisted of a self-paced sentence reading task accompanied by musical sequences from one of three musical grammars: Western, Acquired, and Random. Standard garden-path and semantic anomaly results were found in that those sentences that contained an unexpected linguistic event were more difficult to process as evident in increased reading times at the critical points of the sentences. No interaction between language-syntactic unexpectancy and music-syntactic unexpectancy was found. These results were consistent across all musical grammar conditions.

The results of the first phase of the experiment, where participants were exposed to the artificial grammar, support and extend the findings of Loui & Wessel (2006). The results found in the present study did not replicate those seen in the experiment (Experiment 3) which the present procedure is based upon. However, they do replicate those results found by Loui & Wessel in their Experiment 1 during which participants were exposed to fewer melodies for a shorter amount of time; participants were able to recognize previously heard melodies but were unable to generalize what they had learned to novel melodies.

The difference in results despite using a similar procedure may be due to the population used. Loui and Wessel's participants were musicians (i.e., had over 5 years of musical training) while the majority (89.4%) of the participants in the present study were non-musicians (i.e. had less than 5 years of musical training). Non-musicians may need more exposure time in order to gain enough knowledge of the grammar to generalize to novel melodies. Research has shown the musicians have a better learning and memory for novel melodies than non-musicians (Korenman, 2007). This may be related to changes how auditory information is processed in the brain; that is, musicians show earlier and larger brain responses to auditory information (Musacchia, Sams, Skoe, & Kraus, 2007).

The results of the second phase of the experiment, the sentence-reading task, did not replicate Sleve et al.'s (2009) findings of an interaction between linguistic-syntactic unexpectancy and music-syntactic unexpectancy for the Western condition. The present results were also unable to extend Sieve et al.'s findings to an artificial grammar. However, there were gender differences seen in the sentence-reading task for the Western music condition. Females showed a pattern of RTs consistent with Sleve et al.'s results in that there was an increase in reading time when simultaneous music- and language-syntax errors occurred.

There are a number of reasons why the present study was unable to support the results of Sieve et al (2009). For one, the difference in the number of non-musicians versus musicians differed between the two studies. The present study had only 10% musicians while 50% of Slevc et al.'s participants were considered musicians. Non-musicians may not be as sensitive as musicians to violations of music grammar and therefore may not be as affected by those violations (Koelsch et al., 2003). Musicians are considered to have more specific representations of music-syntactical regularities and therefore have more precise expectations for musical grammar than non-musicians due to their greater exposure to and more explicit training of these tonal relationships (Koelsch, Schmidt, & Kansok, 2002).

Another reason could be the difference in sample size. The present study has 144 participants, which breaks down to 12 in each condition. Sievc et al. had twice as many participants in each condition. Due to the small effects associated with self-paced reading studies, a larger sample size may be needed to find significant results. Also, the difference in study procedures may have affected the results of the present study. Sieve et al's experiment consisted of only the sentence-reading task which lasts approximately 10 minutes. The present study took around 60 minutes to complete due to the addition of the grammar exposure phase and musical tasks. Since the sentence-reading task was the last phase of the present study, participants may have been fatigued after completing the first phase of the experiment. Gender differences were not discussed by Sleve et al. nor are they often discussed in the music cognition literature.

Overall, the present study was able to show that participants were able to learn an artificial grammar but were unable to form a precise representation of the new grammar which would have allowed them to generalize to novel melodies. However, the present study was unable to show an interaction between music and language processing for either the Western or Acquired grammar.

The present results support the future use of this artificial grammar as means to study music acquisition but future research should take into account the abilities of musicians versus non-musicians. Future research should also explore the differences between musicians and non-musicians on tasks of music acquisition and how musical expertise affects the syntactical representations formed during the learning of a new music grammar. At this point in time, studies have either considered one group or the other (Musicians: Loui & Wessel, 2006; Non-musicians: Creel, Newport, & Aslin, 2004; Saffran et al., 1999) and a comparison of the two groups may prove to be useful in the discussion of differences between the two.

However, the way in which participants are categorized as musicians or non-musicians should be reconsidered for future research. At this time individuals are considered a musician if they have received a pre-determined number of years of private musical instruction. Yet, this may not prove to be the most exact method for defining musical expertise. A music expertise scale should be created that is able to test participants' performance on a number of musical tasks in order to quantitatively determine the musicianship of the individual. This would enable researchers to better explore the differences between musicians and non-musicians.

Music acquisition using this grammar should also be explored in other populations including children, bilingual individuals, and those with pervasive developmental disorders and brain damage that affect language learning, comprehension, and production. Studies looking at these populations may prove useful in further discussing the similarities and differences between music and language processing.

Gender's role in music processing should also be further explored. Research has shown differences between males and females in the structure and function of the brain during language processing. Despite the similarities between language and music, there has been relatively little research observing gender differences in music perception and cognition (Koelsch et al., 2003). Gender should be reported and considered in current research in the area of music cognition.

The present study used an artificial music grammar (Loui & Wessel, 2006) in order to test the ability of non-musician participants to learn a new music grammar as well as to observe a possible interaction between music and language syntax processing. Although participants were able to learn the artificial music grammar, a language task was not affected by errors in the new music grammar as has been found with Western music-syntax errors (Sieve, Rosenberg, & Patel, 2009). The results support the possibility of a shared learning mechanism between music and language using statistical probabilities to acquire the syntactical regularities of a grammar. However, the results of the present study were unable to provide evidence for the SSIRH using a self-paced reading task.

References

Akmajian, A., Demers, R., & Harnish, M. (1984). *Linguistics: An Introduction to Language and Communication.* Cambridge, MA: The MIT Press.

Andreou, G., Vlachos, F., & Andreou, E. (2005). Affecting factors in second language learning. *Journal of Psycholinguistic Research*, 34(5), 429-438.

Atchley, R., Rice, M., Betz, S., Kwasny, K., Sereno, J., & Jongman, A. (2006). A comparison of semantic and syntactic event related potentials generated by children and adults. *Brain & Language*, 99, 236-246.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain*, 125, 238-251.

Brown, S., Martinez, M.J., & Parsons, L.M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European Journal of Neuroscience*, 23, 2791-2803.

Burman, D.D., Bitan, T. & Booth, J.R. (2008). Sex differences in neural processing among children. *Neuropsychologia*, 46, 1349-1362.

Chen, E., Gibson, E., & Wolf, F. (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language*, 52, 144-169.

Creel, S.C., Newport, E.L., & Aslin, R.N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5), 1119-1130.

Ditman, T., Holcomb, P., & Kuperberg, G. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*, 44, 927-935.

Eckstein, K. & Friederici, A. (2006). It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience*, 18(10), 1696-1711.

Friederici, A. & Meyer, M. (2004). The brain knows the difference: two types of grammatical violations. *Brain Research*, 1000, 72-77.

Hyde, K. & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, 15(5), 356-360.

Jackendoff, R. (2002). *Foundations of Language.* New York: Oxford University Press.

Janata, P., Birk, J.L., Tillman, B., & Bharucha, J.J. (2003). Online detection of tonal pop-out in modulating contexts. *Music Perception*, 20, 283-305.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, 15, 207-212.

Koelsch, S. (2006). Significance of Broca's area and the ventral premotor cortex for music-syntactic processing. *Cortex*, 42, 518-520.

Koelsch, S., Gunter, T.C., Friederici, A.D., & Schroger, E. (2000). Brain indices of music processing: 'non-musicians' are musical. *Journal of Cognitive Neuroscience*, 12, 520-541.

Koelsch, S., Gunter, T.C., von Cramon, D.Y., Zysset, S., Lohmann, G., & Friederici, A.D. (2002). Bach speaks: A cortical 'language-network' serves the processing of music. *Neuroimage*, 17, 956-966.

Koelsch S., Gunter T.C., Wittforth, M., & Sammler, D. (2005). Interaction between syntax processing in language and music: An ERP study. *Journal of Cognitive Neuroscience*, 17, 1565-1577.

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophusiology*, 44, 476-490

Koelsch, S., Maess, B., Grossmann, T., & Friederici, A.D. (2003). Electric brain responses reveal gender differences in music processing. *NeuroReport*, 14(5), 709-713.

Koelsch, S., Schmidt, B., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, 39, 657-663.

Korenman, L.M. (2007). Individual differences in learning and remembering music: Auditory versus visual presentation. *Journal of Resarch in Music Education*, 55(1), 48-64.

Krumhansl, C. (1990). *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press.

Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research*, 1142, 160-177.

Loui, P. & Wessel, D. (2006). Acquiring new musical grammars – a statistical learning approach. *Paper presented at the 9th International Conference on Music Perception and Cogntion.*

Maess, B., Koelsch, S., Gunter, T.C., & Friederici, A.D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, 4, 540-545.

McKoon, G. & Ratcliff, R. (2007). Interactions of meaning and syntax: implications for models of sentence comprehension. *Journal of Memory and Language*, 56, 270-290.

McMullen, E. & Saffran, J. (2004). Music and language: A developmental comparison. *Music Perception*, 21, 289-311.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *PNAS*, 104(40), 15894-15898.

Patel, A.D. (2008). *Music, Language, and the Brain.* Oxford: Oxford University Press.

Patel, A.D. (2007). Language, music, and the brain: a resource-sharing framework. *Keynote paper for "Language and Music as Cognitive System,"* May 11-13, 2007, Cambridge University, England.

Patel, A.D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Annals New York Academy of Sciences*, 1060, 59-70.

Patel, A.D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6, 674-681.

Patel, A.D. & Iverson, J.R. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology*, 22(7-8), 776-789.

Patel, A.D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10, 717-733.

Peretz, I. & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6, 688-691.

Peretz, I., Kolinsky, R., Tramo, M., Labreque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, 117, 1283-1302.

Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science*, 21, 679-693.

Saffran, J.R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, 44, 493-515.

Saffran, J.R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*,47, 172-196.

Saffran, J.R. (2003). Musical learning and language development. *Annals New York Academy of Science*, 999, 1-5.

Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27-52.

Saffran, J.R., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, 107, 479-500.

Shaywitz, B.A., Shaywitz, S.E., Kenneth, R., Constable, R.T. et al., (1995). Sex differences in the functional organization of the brain for language. *Nature*, 373, 607-609.

Slevc, L.R., Rosenberg, J.C., & Patel A.D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bullitin & Review*, 16(2), 374-381.

Tillmann, B., Janata, P., & Bharucha, J.J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, 16, 145-161.

Truswell, J., Tanenhaus, M., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology*, 19, 528-553.

van Gompel, R., Pickering, M., Pearson, J., & Jacob, G. (2006). The activation of inappropriate analyses in garden-path sentences: evidence from structural priming. *Journal of Memory and Language*, 55, 335-362.

Vincenzi, M., Job, R., Di Matteo, R., Anguilli, A., Penolazzi, B., Ciccarelli, L., & Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language*, 85, 280-296.

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain & Language*, 108, 175-183.

Yamada, Y. & Neville, H. (2007). An ERP study of syntactic processing in English and nonsense sentences. *Brain Research*, 1130, 167-180.

Appendix

Musical Expertise Questionnaire:

1) Are you male or female?: Male / Female / No Response

2) What year are you in school?: Freshman / Sophomore / Junior / Senior / Other

3) Are you a music major?: Yes / No

4) Do you play an instrument or sing?: Yes / No

5) Can you read music notation?: Yes / No

6) Have you received private music instruction?: Yes / No

7) If so, how many years of instruction have you had?: 0-1 year / 2-3 years / 4-5 years / 6-7 years / 8 or more years

8) Do you consider yourself a musician?: Yes / No

9) Do you have perfect pitch (a.k.a. absolute pitch)?: Yes / No