12-2011

The role of metrical structure in tonal knowledge acquisition

Matthew Rosenthal
University of Nevada, Las Vegas

THE ROLE OF METRICAL STRUCTURE IN TONAL KNOWLEDGE ACQUISITION

By

Matthew Rosenthal

Bachelor of Science
Indiana University
2007

A thesis submitted in partial fulfillment of
the requirements for the

Master of Arts in Psychology
Department of Psychology
College of Liberal Arts

Graduate College
University of Nevada, Las Vegas
December 2011

THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Matthew Rosenthal

entitled

The Role of Metrical Structure in Tonal Knowledge Acquisition

be accepted in partial fulfillment of the requirements for the degree of

Master of Arts in Psychology
Department of Psychology

Erin Hannon, Committee Chair

Joel Snyder, Committee Member

Mark Ashcraft, Committee Member

Gabrielle Wulf, Graduate College Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

December 2011

ABSTRACT

THE ROLE OF METRICAL STRUCTURE IN TONAL KNOWLEDGE ACQUISITION

By

Matthew Rosenthal

Erin Hannon, Examination Committee Chair
Assistant Professor of Psychology
University of Nevada, Las Vegas

Experienced listeners possess a working knowledge of pitch structure in Western music, such as scale, key, harmony, and tonality, which develops gradually throughout childhood. It is commonly assumed that tonal representations are acquired through exposure to the statistics of music, but few studies have investigated potential learning mechanisms directly. In Western tonal music, tonally stable pitches not only have a higher overall frequency of occurrence, but they may occur more frequently at strong than weak metrical positions, providing two potential avenues for tonal learning. Two experiments employed an artificial grammar learning paradigm to examine tonal learning mechanisms. During a familiarization phase, we exposed nonmusician adult listeners to a long (whole tone scale) sequence with certain distributional properties. In a subsequent test phase we examined listeners' learning using grammaticality or probe tone judgments. In the grammaticality task, participants indicated which of two short test sequences conformed to the familiarization sequence. In the probe tone task, participants provided fit ratings for individual probe tones following short "reminder" sequences. Experiment 1 examined learning from overall frequency of occurrence. Grammaticality judgments were significantly above chance (Exp. 1a), and probe tone ratings were predicted by frequency of occurrence (Exp. 1b). In Experiment 2 we presented a familiarization sequence containing one sub-set of pitches that occurred more frequently on strong than on weak metrical positions and another sub-set that

did the opposite. Overall frequency of occurrence was balanced between both sub-sets. Grammaticality judgments were again above chance (Exp. 2a) and probe tone ratings were higher for pitches occurring on strong metrical positions (Exp. 2b). These findings implicate metrical structure in tonal knowledge acquisition.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Most Western music is composed so that pitches are structured with respect to a central reference pitch, or tonic, according to the principles of musical tonality (Krumhansl, 1990). Theorists have suggested that tonality is beneficial to listeners because it facilitates the formation of musical expectations, which listeners perceive as changes in stability (Huron, 2006; Lerdahl & Jackendoff, 1996). Naturally pleasing or consonant sounds have typically been considered to play a fundamental role in determining the pitch stability relationships of tonality, with more stable relationships occurring between pitches that sound consonant and are related by simple ratios (James, 1993; Schellenberg & Trehub, 1996; Large, 2010). Tonality even has been described as an innate Universal Grammar (Bernstein, 1973). Natural properties of sound such as ratio simplicity likely play an important role in tonality as simple ratios are common in the scales of various musical cultures (Burns, 1999). However, natural sound properties are unlikely to account for all aspects of tonality as experience with music is critical in acquiring knowledge of the musical structures in a given culture (Bigand & Poulin-Charronnat, 2006; Hannon & Trehub, 2005a; Hannon & Trehub, 2005b; Krumhansl & Keil, 1982; Lynch, Eilers, Oller, & Urbano, 1990; Trainor & Trehub, 1992; Trainor & Trehub, 1994; Trehub & Hannon, 2006). Previous studies have shown that listeners could acquire tonal knowledge by internalizing statistical regularities in music, such as the frequency with which individual pitches occur (Loui, Wessel, & Hudson Kam, 2010), transitional probabilities between adjacent pitches (Saffran, Johnson, Aslin, & Newport, 1999), transitional probabilities between adjacent chords (Jonaitis, & Saffran, 2009), and in certain situations, non adjacent transitional probabilities (Creel, Newport, & Aslin, 2004; Endress, 2010). However, previous research on tonal knowledge acquisition has not explored potential contributions of temporal structure despite known interactions between temporal and pitch structure during music listening (Ellis & Jones, 2009; Hannon & Johnson, 2005; Hannon, Snyder, Eerola, & Krumhansl, 2004; Krumhansl, 2000). Here we show that listeners are sensitive to statistical contingencies between metrical and pitch structure and argue that this sensitivity contributes to tonal knowledge acquisition.

If tonal knowledge is acquired through exposure to music, then strength of tonal knowledge should increase with age and experience. Accordingly, developmental investigations have demonstrated that adult-like knowledge of Western tonality emerges gradually during childhood. For instance, when monitoring a repeating melody for pitch changes, eight-month-old infants notice in key and out of key changes equally well, whereas adult listeners are better at detecting out of key changes (Trainor & Trehub, 1992). These findings indicate that very young listeners do not posses fundamental knowledge of key membership. Listeners begin to demonstrate tonal knowledge in early childhood, with five year olds distinguishing pitches by key membership and seven-year olds and adults distinguishing pitches by key and chord membership (Trainor & Trehub, 1994). A similar trajectory in tonal knowledge acquisition is demonstrated by participants' ratings of melody quality (good or bad). When children (first through sixth grade) and adults are asked to rate simple melodies ending on different pitches, all ages tended to provide higher goodness ratings for melodies ending on in-key pitches and lower goodness ratings for melodies ending on out-of-key pitches (Krumhansl & Keil, 1982). However, it is not until third grade that children distinguish melodies based on the ending pitch's membership in the tonic triad and not until adulthood that listeners distinguish the tonic from other tonic triad members (Krumhansl & Keil, 1982). Together, the above studies describe the developmental trajectory of tonal knowledge acquisition, which spans infancy to adulthood, presumably as listeners gain experience listening to Western tonal music.

The statistical regularities in music provide one potential means through which listeners could acquire tonal knowledge (Huron, 2006; Krumhansl, 1990). Krumhansl & Kessler (1982) investigated the perceived stability of pitches in the Western major scale. Participants were presented with a short context (a single chord, a melodic sequence, or a three-chord sequence) intended to induce major tonality. After the context, participants were asked to indicate how well a probe tone, any one of the twelve chromatic pitches, fit with the context sequence they just heard. Participants' ratings indicated a hierarchy of stability relative to the most highly rated tonic pitch, matching music theoretic predictions. Importantly, the resulting profile of ratings was strongly correlated with the frequency of occurrence of sounded pitches in actual music (Knopoff & Hutchinson, 1983; Krumhansl, 1990; Youngblood, 1958). Thus, tonal stability could reflect the internalization of musical statistics, particularly pitch frequency of occurrence (Krumhansl, 1990). Recent empirical investigations have supported this claim. Loui, Wessel, & Hudson Kam (2010) asked participants to provide probe tone ratings for short melodic sequences created using pitches from an unfamiliar musical scale. Probe tone ratings were positively correlated with frequency of occurrence. Creel & Newport (2002) performed a similar experiment using a different unfamiliar (whole tone) scale. Once again, participants provided the highest probe tone ratings for pitches that occurred frequently in context sequences.

Listeners also appear to acquire tonal knowledge by internalizing the statistical structure of sequential relationships between pitches. For instance, in Western tonal music, a given pitch tends to predict a small set of possible next pitches (Huron, 2006). A listener potentially could acquire knowledge of the transitional relationships between pitches and use it to infer tonal prominence. In behavioral tasks, when asked to judge the similarity of a stable pitch and unstable pitch following a tonal context, participants respond that the two pitches are more similar when the unstable pitch precedes the stable one, than vice versa (Krumhansl, 1979). This trend in participants' judgments could reflect the tendency in Western music for unstable pitches to more accurately predict the onset of stable pitches than vice versa (Huron, 2006). Accordingly, several studies have shown that listeners are sensitive to the sequential structure governing event-to-event transitions in musical sequences (Endress, 2010; Jonaitis, & Saffran, 2009; Saffran et al., 1999), implicating sequential relationships in the perception of pitch stability.

In the speech domain, the ability to learn about predictive relationships is thought to underlie acquisition of linguistic syntactical structure (Saffran, 2001). Unlike event-to-event sequential structure, predictive relationships in linguistic syntax often occur between elements that aren't directly adjacent to each other. Typically participants are challenged by non-adjacent statistical relationships, only detecting them in specific situations, such as when non-adjacent elements are highlighted by pitch changes (Creel et al., 2004) or when non-adjacent relationships are more reliable than adjacent relationships (Gomez, 2002). One means by which non-adjacent elements are accented in speech is through syllable stress (Cutler & Butterfield, 1992; Jusczyk, 1999). Syllable stress is thought to facilitate syntax acquisition through the occurrence of rhythmic accents at major syntactic boundaries, allowing learners to segment the speech stream into meaningful syntactical units (Höhle, 2009). Similarly to syllable stress in speech, metrical structure (meter) in music organizes non adjacent sounds on the basis of accents. Specifically, meter is defined as the periodic, hierarchically organized alternation of strong and weak accents that is inferred from the occurrence of periodic "phenomenal" accents (e.g. changes in duration, volume, contour, etc.) on the musical surface (Lerdahl & Jackendoff, 1996). When listeners detect periodicity in a sequence of phenomenal accents, a metrical representation presumably is activated. Following this activation, listeners interpret musical events with respect to a pattern of nested hierarchical levels of periodicity, some levels occurring at faster timescales (e.g. two or three times as frequently as the main periodicity or *tactus*) and some levels occurring at slower timescales (e.g. half as frequently). The points at which faster and slower levels of periodicity coincide are perceived as accented or metrically strong, while points in which fewer periodicities overlap are perceived as metrically weak. In principle, meter could highlight events in a pitch sequence, providing another means by which listeners could infer tonal prominence.

Some evidence suggests that pitch and temporal structure are processed independently of each other in early stages of processing, and that any perceptual interactions between the two dimensions occurs only at late stages (Peretz & Kolinsky, 1993). Potentially, this view could explain previous observations that pitch and temporal structure are perceptually independent. In Palmer and Krumhansl (1987), participants were asked to judge the completeness of melodies whose ending notes varied in tonal and metrical stability. It was assumed that if meter and pitch are integrated in processing, then for melodies ending on stable pitches and on strong metrical positions, participants should provide higher ratings than would be predicted by adding the individual contributions of tonal and metrical stability. The best fitting model simply added the two components, indicating that participants were not, by the authors' definition, integrating meter and pitch. If meter and pitch are processed independently, we might predict that meter would not play a role in the acquisition of musical pitch structure. However, even if pitch and temporal structure are independent at an early stage of processing, this does not rule out the possibility that metrical information could provide structure to pitch input and thereby change how listeners infer tonal prominence from pitch patterns. Consistent with this possibility, listeners appear to associate tonal stability and metrical strength. For example, in jazz improvisation, musicians tend to play tonally stable pitches on metrically strong positions and tonally unstable pitches on metrically weak positions (Järvinen, 1995), providing an example of correlated pitch-time structure in spontaneous musical output. Such pitch-meter correlations might also be internalized by listeners. When participants are asked to indicate whether a probe tone occurred on or off the beat in short melodies, they are biased to report that tonally stable probe tones occurred on the beat (i.e. on strong metrical positions) and tonally unstable probe tones occurred off the beat (i.e. on weak metrical positions), even though probe tones were equally likely to occur on any metrical position (Prince, Thompson, & Schmuckler, 2009). Thus, tonal stability potentially reflects acquired knowledge of pitch-meter correlations in actual music.

Young learners appear to posses knowledge of meter sophisticated enough to facilitate tonal knowledge acquisition. Hannon & Johnson (2005) investigated seven-month-old infants' discrimination of melodies based on the frequency with which individual pitches occurred on strong or weak metrical positions. Participants were habituated to short triple-meter melodies in which one subset of pitches tended to occur on strong metrical positions and another subset tended to occur on weak metrical positions. In the test phase, participants discriminated novel melodies on the basis of the melodies' conformity to the distribution of strong and weak pitches during habituation. This finding indicates that infants associate specific pitches with strong and weak metrical positions. However, it does not necessarily demonstrate that infants can use this information to infer tonal prominence. If meter contributes to tonal knowledge acquisition, then metrically strong pitches should be perceived as more tonally prominent than metrically weak pitches.

The experiments in this thesis simulate the learning challenge faced by listeners who are confronted with novel, complex musical structures in the environment. In an attempt to replicate previous findings (Creel & Newport, 2002; Loui et al., 2010), Experiment 1 examines whether listeners can use individual pitch frequency of occurrence to learn about pitch structure in a novel musical context. Participants are first exposed to a short familiarization sequence in which individual pitches vary in their frequency of occurrence. Participants are asked to provide a grammaticality judgment by choosing which of two short test sequences is consistent with the familiarization (Experiment 1a), or they are asked to rate how well a probe tone fits with a preceding "reminder" sequence that mirrors the statistical structure of the familiarization sequence (Experiment 1b). If participants are sensitive to the distributional information in the familiarization sequence, they should readily discriminate between the consistent and inconsistent test melodies in the grammaticality judgment task. If participants not only pick up on distributional information but also use that information to infer tonal prominence, then more frequently occurring pitches should receive higher probe tone ratings. Experiment 2 extends the basic method of Experiment 1 to examine whether listeners can use metrical position to learn about pitch structure. In this experiment, rhythmic familiarization sequences contain a subset of pitches that occurs more often on strong metrical positions and a separate subset that occurs more frequently on weak metrical positions. Crucially, individual pitches in both subsets have the same overall frequency of occurrence, so that metrical position is the only cue listeners can use to learn about pitch structure. As in Experiment 1, Experiment 2a requires participants to choose which of two test sequences is consistent with a familiarization sequence and Experiment 2b requires them to give probe tone ratings for individual pitches following short context sequences. If adults are sensitive to contingencies between pitch and meter, as infants are

(Hannon & Johnson, 2005), they should be able to discriminate consistent from inconsistent test sequences. If participants are also able to use metrical position to infer tonal prominence, then probe tone ratings should be higher for pitches occurring on strong metrical positions and lower for pitches occurring on weak metrical positions.

CHAPTER 2

EXPERIMENT 1A

Method

Participants

Forty-one undergraduates taking Psychology courses at the University of Nevada Las Vegas participated in the study for course credit. One participant's data were excluded because of a technical difficulty during the experiment. Participants (17 female, mean age = 20.6; age range = 18-39) were randomly assigned to one of two familiarization conditions (A or B), with 20 participants in each condition. The average duration of formal music instruction (according to questionnaire responses) was 1.8 years (range = 0-19 years). Participants had normal hearing and no history of hearing difficulties.

Apparatus

Participants sat at an Apple Mac Mini computer. Stimuli were presented using Sony noise canceling headphones. The experiment was presented and controlled using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993), which displays instructions, plays stimuli, and records participants' keyboard presses.

Stimuli

Two different familiarization sequences (Familiarization A and Familiarization B) were created. All sequences were created using a MIDI sequencer (Digital Performer 4.6) and converted to .wav format using the software sampler Reason 3.0 “bright piano” timbre. Each familiarization sequence contained 480 events (i.e. sounded pitches) which were 250 ms in duration and separated by a 250 ms inter-onset interval. Based on visual inspection of the stimulus waveform, rise times were approximately 10 ms and fall times were approximately 240 ms. Each event was assigned one of six pitches from the whole tone scale in C (C3-D3-E3-F#3-G#3-A#3). We used the whole tone scale because we wanted to minimize potential biases towards familiar scale structures and Western tonality. The whole tone scale also lacks the perceptually prominent Perfect 5th interval and is rare in popular music. Thus, we did not expect it to strongly activate Western tonal biases (Creel et al., 2002). Each of the 6 pitches of the whole tone scale was assigned to one of three frequency of occurrence conditions: infrequent, moderately frequent, and highly frequent. These conditions determined how often a pitch would occur in a familiarization sequence. For both familiarizations, three infrequent pitches occurred 9% of the time, two moderately frequent pitches occurred 18% of the time, and one highly frequent pitch occurred 36% of the time (see Figure 1). For each pitch, its frequency of occurrence was reversed in Familiarization A and Familiarization B. Thus, if a given pitch was highly frequent in Familiarization A, that same pitch would be infrequent in Familiarization B.

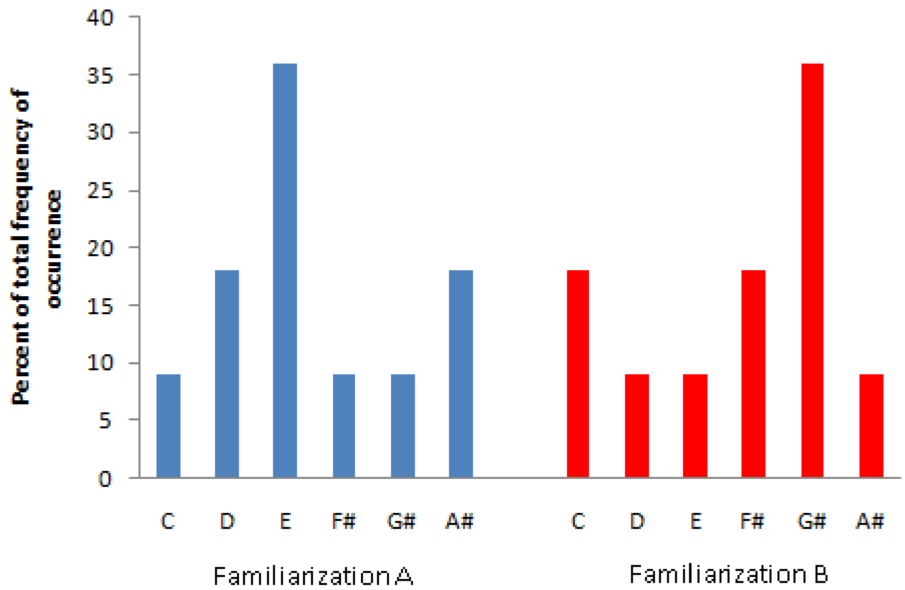

Figure1: Pitch frequency of occurrence in Familiarizations A and B in Experiment 1.

Twelve test sequences were created in the same manner as familiarization sequences, but test sequences were shorter, with 11 events for a total duration of approximately 3 s each. Six test sequences matched Familiarization A (Test A sequences) and six matched Familiarization B (Test B sequences).

Procedure

After signing a consent form and filling out a brief questionnaire, participants were seated at a computer and told that they would hear two minutes of a musical language (either Familiarization A or Familiarization B). They were also told that after the two minutes, they would hear short musical sequences and make judgments about them. During the test phase, which immediately followed the familiarization phase, participants heard two short melodies (one Test A and one Test B) and were asked to indicate whether the first short melody or the second short melody came from the musical language they heard earlier by pressing 1 or 2. Each of four Test A sequences was paired with each of four Test B sequence for a total of 16 Test A-B pairs. Each pair was presented to each participant twice, with the order of presentation counterbalanced so that half the time the matching test stimulus was presented first and half the time it was presented second, totaling 32 test trials.

Results

The percentage of trials in which a participant chose the Test A sequence was computed for each participant. These percentages were submitted to a one way (Familiarization Type [A, B]) ANOVA. The ANOVA revealed a significant effect of Familiarization Type, $F(1,38) = 13.306$, $p < .01$, indicating that participants chose the Test A sequence more after hearing Familiarization A than after hearing Familiarization B, and vice versa (see Figure 2). For subsequent analyses, Test A choices were counted as correct if participants heard Familiarization A, and Test B choices were counted as correct if participants heard Familiarization B. One sample *t*-tests against chance level (50%) indicated that participants were significantly more accurate than chance whether they were presented with Familiarization A, $t(19) = 2.8$, $p < .05$, or Familiarization B, $t(19) = 2.4$, $p < .05$. Percent correct scores were not significantly different between the two groups, $t(38) = .265$, $p = .792$.

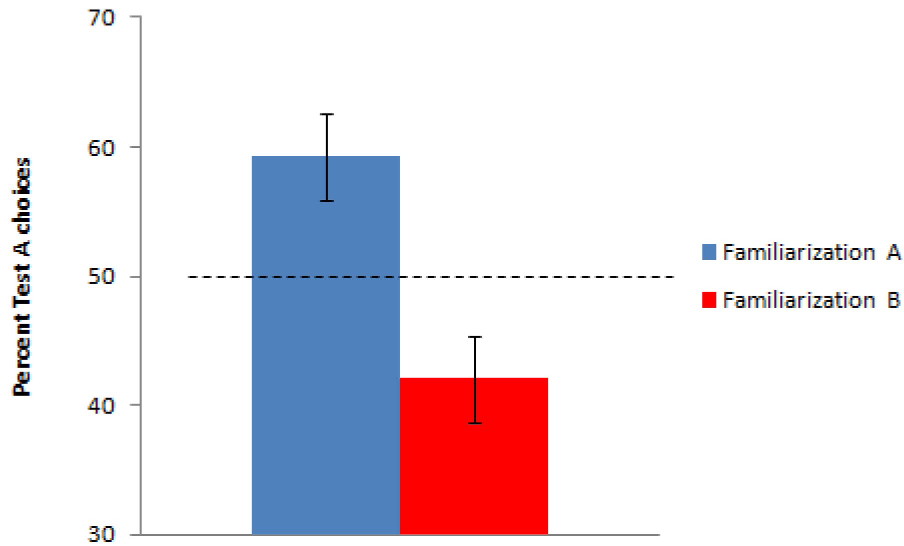

Figure 2: Percentage of Test A choices for Familiarizations A and B in Experiment 1a. Dashed line shows chance performance.

CHAPTER 3

EXPERIMENT 1B

Method

Participants

Forty-four undergraduates from the University of Nevada Las Vegas participated in the study for course credit. Four participants' data were excluded because they did not follow instructions to use all values of the response scale (i.e. they answered with only ones and/or sevens). Participants (15 females; mean age = 19.8; age range = 18-25) were randomly assigned to one of two familiarization conditions (A or B), with 20 participants in each condition. The average duration of formal music instruction was 2.1 years (range = 0-11 years). Participants had normal hearing and no history of hearing difficulties.

Apparatus and Stimuli

The same familiarization sequences were used as in Experiment 1a. On test trials, one of the six matching test sequences was presented, followed by an individual probe tone. Probe tones were one of nine pitches (B2-C3-C#3-D3-E3-F3-F#3-G#3-A#3). Six of the probe tones occurred during the familiarization and test sequences and three probe tones never occurred. Thus, each probe tone was 1 of 4 types: non-occurring, infrequent, moderately frequent, and highly frequent. Each probe tone's timbre, attack time, decay time, and duration matched that of events in the familiarization and test sequences. A 1 s interval separated the end of each test sequence and the presentation of the probe tone. Each new trial began as soon as the participant input a response for the preceding trial.

Procedure

After signing a consent form and filling out a brief questionnaire, participants were seated at a computer and told that they would hear two minutes of a musical language (either Familiarization A or Familiarization B). They were also told that after the two minutes, they would hear short musical sequences and make judgments about them. During the test phase, which immediately followed the familiarization phase, participants heard a test sequence followed by a single probe tone and were asked to rate how well the probe tone fit with the previously heard sequence on a scale of 1-7. Participants were told that 1 meant poor fit and 7 meant good fit. Participants were instructed to use the entire scale. Each of six test sequences was paired with 9 probe tones for a total of 54 trials. Test sequences and probe tones were presented in a random order, but each participant heard the same test sequence-probe tone pairs.

Results

Mean probe tone ratings were averaged for each of the four probe tone conditions: non-occurring, infrequent, moderately frequent, and highly frequent. Probe tone ratings were submitted to a 4 × 2 (Frequency of Occurrence [non-occurring, infrequent, moderately frequent, highly frequent] × Familiarization Type [A, B]) mixed design ANOVA. The ANOVA revealed a significant main effect of Frequency of Occurrence $F(1,38) = 67.146$, $p < .001$, $\eta^2 = .639$. The interaction between Frequency of Occurrence and Familiarization Type was not significant $F(1,38) = 1.249$, $p = .295$, $\eta^2 = .032$. Planned paired samples *t*-tests revealed that highly frequent pitches received higher ratings than moderately frequent pitches $t(39) = 3.452$, p < .01, moderately frequent received higher ratings than infrequent $t(39) = 3.442$, p < .01, and infrequent pitches received higher ratings than non-occurring $t(39) = 8.015$, (see Figure 3).

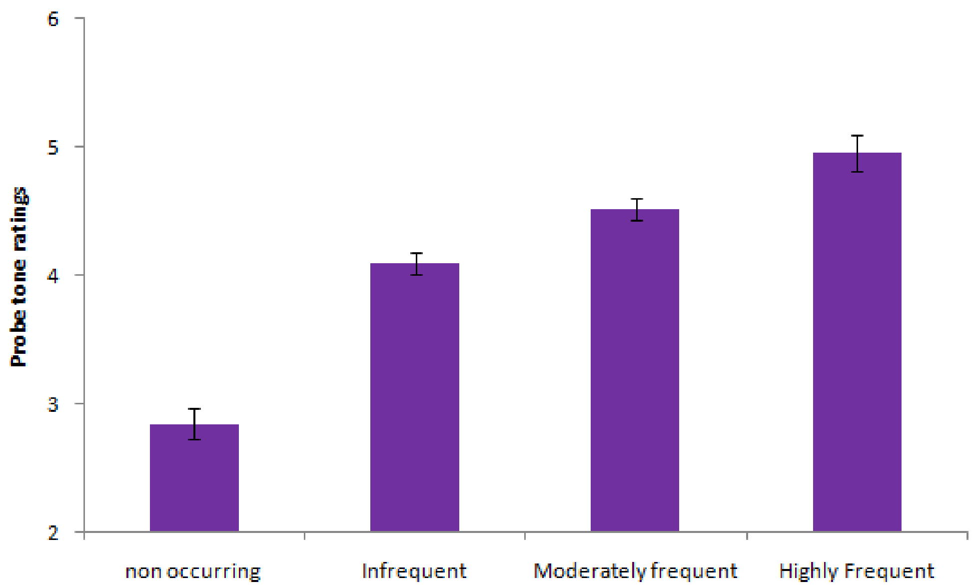

Figure 3: Probe tone ratings as a function of frequency of occurrence in Experiment 1b. Error bars show within-subjects error (Morey, 2008)

Because pitches were pseudo-randomly assigned to each event in the each test sequence, pitches that occurred more often also were more likely to be the final pitch in a context test sequence. Given that recency effects have been shown to influence probe tone ratings (Creel et al., 2002), we wanted to rule out the possibility that our results were driven exclusively by recency effects. We therefore conducted a replication analysis in which we excluded any trial during which the probe tone matched the final note of the test sequence. We again observed a main effect of Frequency of Occurrence, $F(1,38) = 56.97$, $p < .001$, $\eta^2 = .600$. The interaction between Frequency of Occurrence and Familiarization Type was not significant $F(1,38) = 1.435$, $p = .236$, $\eta^2 = .036$. Planned paired samples *t* tests on the recalculated values revealed the same trend of significant differences as was observed in the previous analysis, with highly frequent pitches receiving higher ratings than moderately frequent $t(39) = 3.826$, $p < .001$, moderately frequent receiving higher ratings than infrequent $t(39) = 2.472$, $p < .05$, infrequent pitches receiving higher ratings than non-occurring $t(39) = 7.442$, $p < .001$ (see Figure 4).

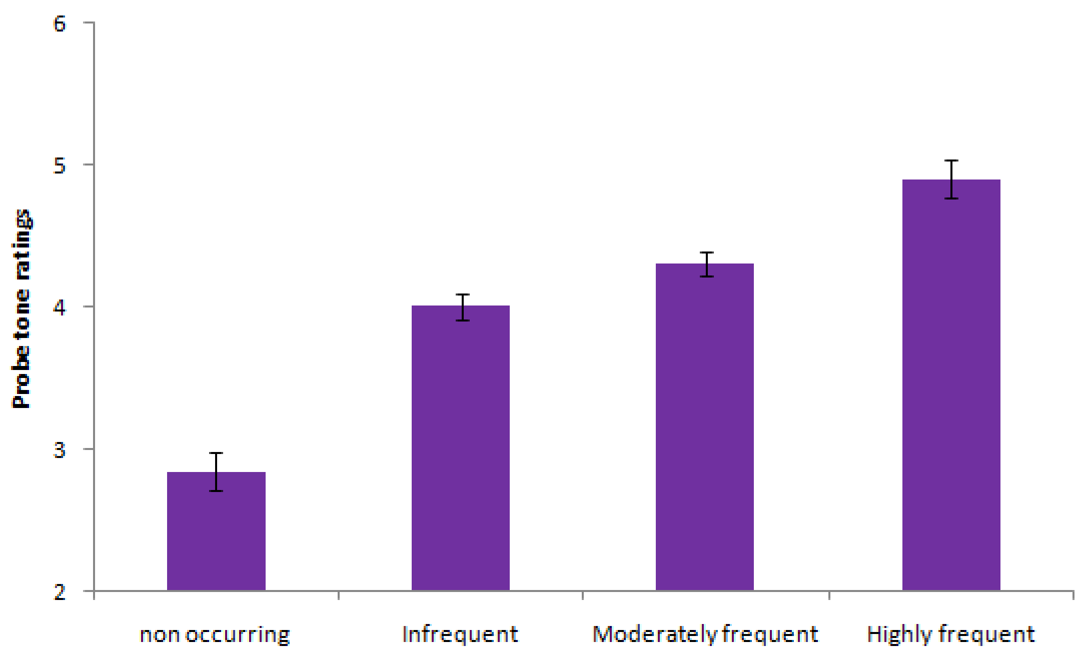

Figure 4: Probe tone ratings as a function of frequency of occurrence excluding trials in which the probe tone was the same as the ending pitch of the context sequence. Error bars show within-subjects error (Morey, 2008)

CHAPTER 4

EXPERIMENT 2A

Method

Participants

Forty undergraduates from the University of Nevada Las Vegas participated in the study for course credit. Participants (26 female; mean age = 19.9; age range = 18-40) were randomly assigned to one of two familiarization conditions (A or B), with n = 20 in each condition. The average duration of formal music instruction was 1.9 years (range = 0-12 years). Participants had normal hearing and no history of hearing difficulties.

Apparatus and Stimuli

Triple-meter rhythms were created based on prior studies (Hannon & Johnson, 2005). For these rhythms, events and temporal group accents occurred most frequently every third position (i.e. every 750 ms). To accomplish this, every third temporal position in the sequence was designated "strong" and all other positions were designated "weak", and silence was assigned to temporal positions according to the following rules adapted from Povel & Essens (1985):

1. No silence occurred on strong metrical positions
2. Events on strong metrical positions were not both preceded and followed by other events
3. Events on weak metrical positions were not followed by silence

Two different familiarization sequences (Familiarization A and Familiarization B) were created using a MIDI sequencer (Digital Performer 4.6). Each familiarization sequence contained 481 temporal units of 250 ms duration. Each temporal unit contained either an event or silence. Events' duration, rise time, and fall time were the same as described in Experiment 1.

Events were comprised of one of six pitches from the whole tone scale in C (C3-D3-E3-F#3-G#3-A#3). Each of the 6 pitches was assigned to one of two classes: strong or weak. These two classes determined whether a pitch tended to occur on strong or weak metrical positions. Thus, the pitch of an event in each familiarization sequence was determined pseudorandomly, with strong and weak pitches occurring approximately 90% of the time on their respective metrical position and approximately 10% of the time on the other position. For each familiarization condition, 3 pitches were strong and 3 pitches were weak. Each pitch's metrical class was reversed between the two familiarizations so that it was strong in one and weak in the other. Crucially, between the two familiarization conditions, a pitch occurred with the same frequency of occurrence regardless of whether it was designated as strong or weak. Statistical distributions for Familiarizations A and B in Experiment 2 are displayed in Figure 5.

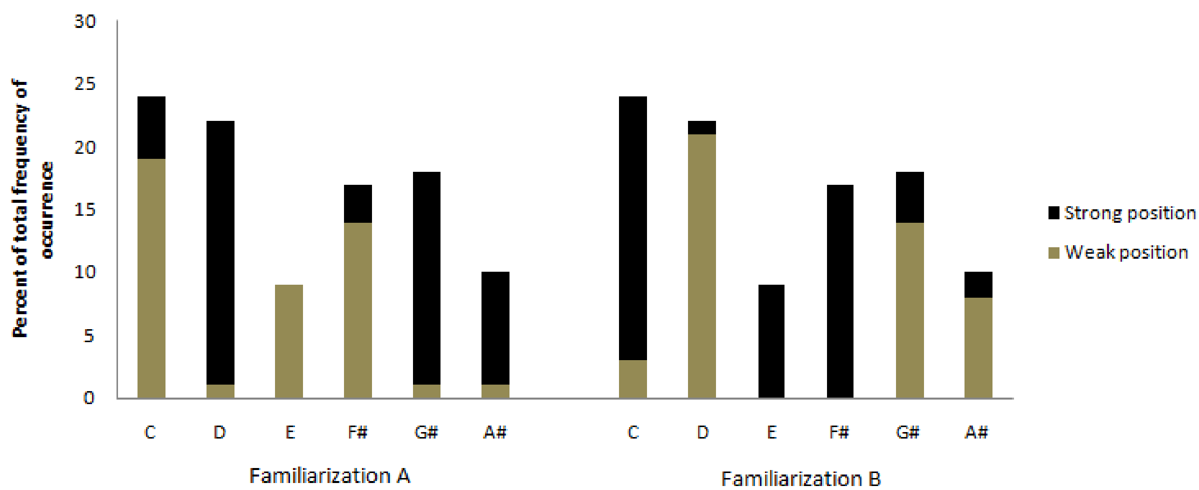

Figure 5: Pitch frequency of occurrence on strong and weak metrical positions in Familiarizations A and B in Experiment 2. Each pitch in one familiarization occurred with the same frequency of occurrence as the same pitch in the other familiarization.

The same rules were used to create twelve short 15-event (out of 24 temporal positions) sequences. Test A sequences matched Familiarization A and Test B sequences matched

Familiarization B. All familiarization and test stimuli started with four measures of a triple meter drum lead-in to make sure the meter was always obvious to the participant.

Procedure

After signing a consent form and filling out a brief questionnaire, participants were seated at a computer and told that they would hear two minutes of a musical language (either Familiarization A or Familiarization B). They were also told that after the two minutes, they would hear short musical sequences and make judgments about them. During the test phase, which immediately followed the familiarization phase, participants heard two short melodies (one Test A and one Test B) and were asked to indicate whether the first short melody or the second short melody came from the musical language they heard earlier by pressing 1 or 2. Each of four Test A melodies was paired with each of four Test B melodies for a total of 16 A-B short melody pairs. Each pair was presented to participants twice, with the order of presentation counterbalanced so that half the time the matching test stimulus was presented first and half the time it was presented second, totaling 32 trials.

Results

The percentage of trials in which a participant chose the Test A sequence was computed for each participant. These percentages were submitted to a one way (Familiarization Type [A, B]) ANOVA. The ANOVA revealed a significant effect of Familiarization Type, $F(1,38) = 18.909$, $p < .001$, indicating that participants chose the Test A sequence more after hearing Familiarization A than after Familiarization B, and vice versa for Test B sequences (see Figure 6). For subsequent analyses, Test A choices were counted as correct for participants hearing Familiarization A, and Test B choices were counted as correct for participants hearing

Familiarization B. One sample *t*-tests against chance level (50%) indicated that participants were significantly more accurate than chance whether they were presented with Familiarization A, $t(19) = 2.53$, $p < .05$, or Familiarization B, $t(19) = 3.57$, $p < .01$. Percent accuracy was not significantly different between the two groups, $t(38) = 1.026$, $p = .311$.

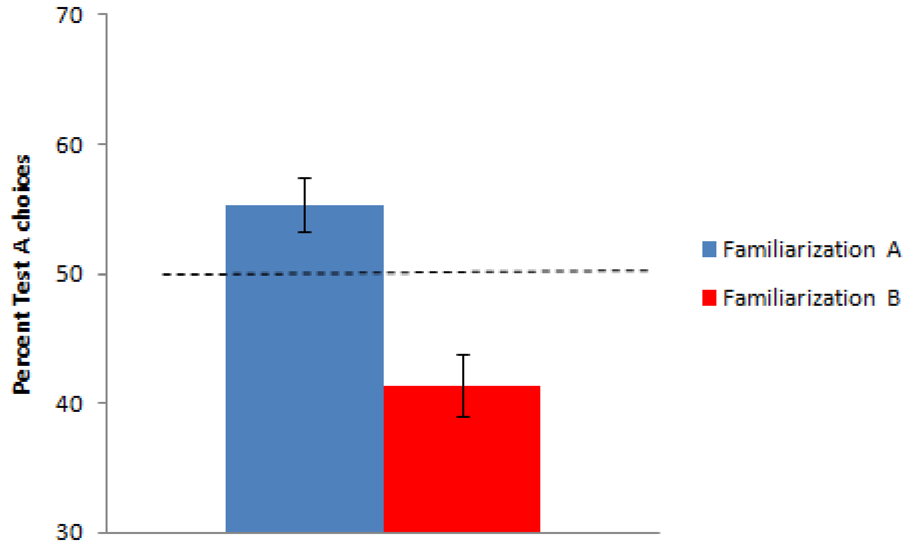

Figure 6: Percentage of Test A choices for Familiarizations A and B in Experiment 2a. Dashed line shows chance performance.

CHAPTER 5

EXPERIMENT 2B

Method

<u>Participants</u>

Forty-two undergraduates from the University of Nevada Las Vegas participated in the study for course credit. Two participants' data were excluded because they did not use intermediate values of the response scale. Participants (23 female; age range = 18-24) were randomly assigned to one of two familiarization conditions (A or B), with n = 20 in each condition. The average amount of music training was 2.7 years (range = 0-11 years). Participants had normal hearing and no history of hearing difficulties.

Apparatus and Stimuli

The same familiarization sequences were used as in Experiment 2a. On test trials, one of six test sequences that matched the familiarization was presented, each followed by individual probe tones. Probe tones were one of nine pitches (B2-C3-C#3-D3-E3-F3-F#3-G#3-A#3). Six of the probe tones occurred during the familiarization and test sequences and three never occurred. Thus each probe tone was 1 of 3 types: strong, weak, or non occurring. Each probe tone's timbre, attack time, decay time, and duration matched that of events in the familiarization and test sequences. A 1 s interval separated the end of each test sequence and the presentation of the probe tone. Each new trial began as soon as the participant input a response for the preceding trial.

Procedure

After signing a consent form and filling out a brief questionnaire, participants were seated at a computer and told that they would hear two minutes of a musical language (either Familiarization A or Familiarization B). They were also told that after the two minutes, they would hear short musical sequences and make judgments about them. During the test phase, which immediately followed the familiarization phase, participants heard a test sequence followed by a single probe tone and were asked to rate how well the probe tone fit with the previously heard sequence on a scale of 1-7. Participants were told that 1 meant poor fit and 7 meant good fit. Participants were instructed to use the entire scale. Participants' data were excluded if they answered with a 1 or 7 on every trial. Each of six test sequences was paired with 9 probe tones for a total of 54 trials. Test sequences and probe tones were presented in a random order, but each participant heard the same test sequence-probe tone pairs.

Results

Mean probe tone ratings were averaged for metrically strong and metrically weak probe tones and submitted to a 2 × 2 (Metrical Strength [strong, weak] × Familiarization Type [A, B]) mixed design ANOVA. The ANOVA revealed a main effect of Metrical Strength, $F(1,38) = 23.49$, $p < .001$, $\eta^2 = .382$ (see Figure 7), with strong pitches receiving higher fit ratings than weak pitches. There was also an interaction between Metrical Strength and Familiarization Type, $F(1,38)=4.297$, $p<.05$, $\eta^2 = .102$. Paired samples *t* tests confirmed that strong pitches were rated higher than weak pitches whether participants were presented with Familiarization A $t(19) = 4.390$, $p < .001$, or Familiarization B $t(19) = 2.254$, $p < .05$ (see Figure 8).

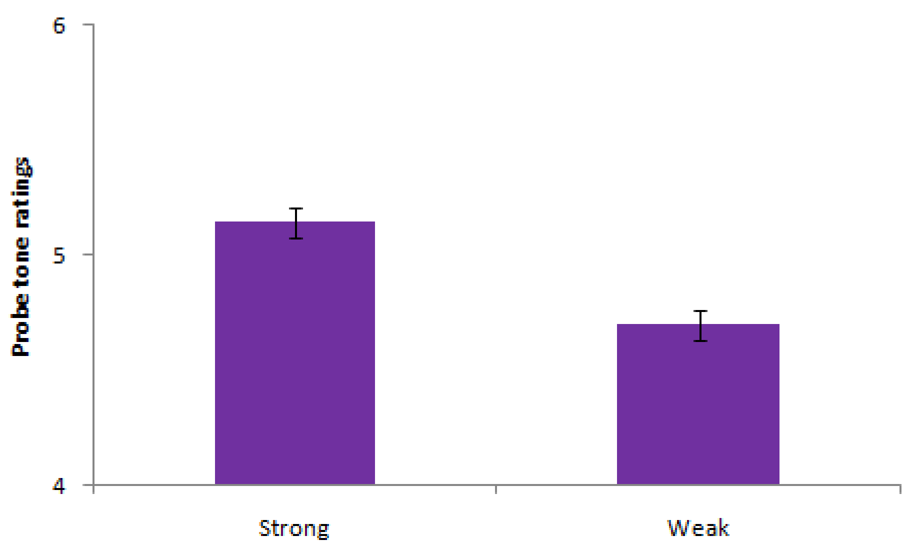

Figure 7: Probe tone ratings for strong and weak pitches in Experiment 2b. Error bars show within-subjects error (Morey, 2008)

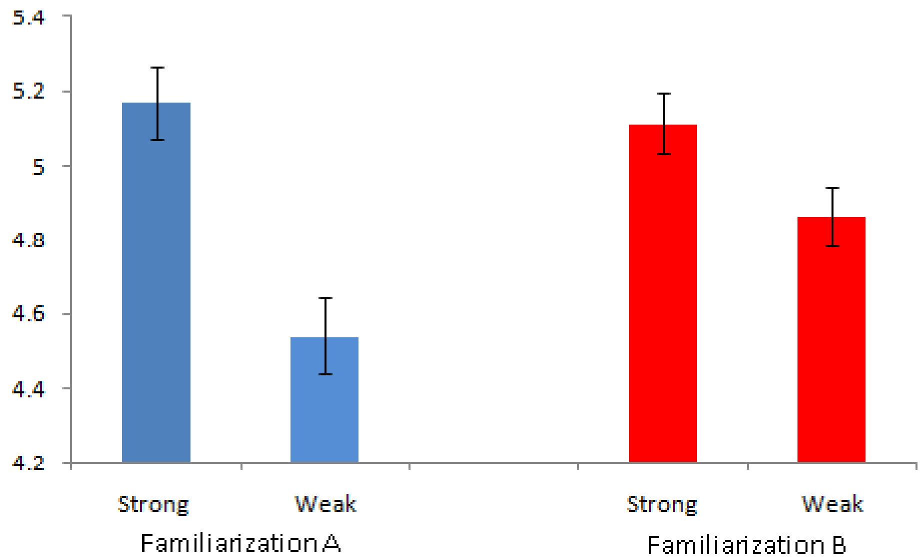

Figure 8: Probe tone ratings for strong and weak pitches in Familiarizations A and B in Experiment 2b. Error bars show within-subjects error (Morey, 2008)

CHAPTER 6

GENERAL DISCUSSION

In Experiment 1, participants successfully matched test sequences based on the frequency of occurrence of sounded pitches and provided higher probe tone ratings for frequent pitches. These findings replicate prior work (Loui et al., 2010; Creel et al. (2002) and support the notion that tonal knowledge is acquired through statistical regularities in music, particularly frequency of pitch occurrence (Krumhansl, 1990). Prior work documented an association between probe tone ratings for individual pitches and their frequency of occurrence distributions in composed music, but Experiment 1, like the studies it replicates (Creel et al., 2002; Loui et al., 2010), provides a causal demonstration that frequency of occurrence can influence listeners' perception of tonal prominence after only a few minutes of listening. Experiment 2 provides novel evidence that musical meter can also influence pitch learning, as participants successfully matched test stimuli to a familiarization sequence based on pitch-meter contingencies and provided higher probe tone ratings for pitches occurring on strong metrical positions.

The present experiments thus provide robust evidence that both frequency of occurrence information and metrical strength can influence learning in an artificial grammar learning paradigm. This statistical information could therefore in principle contribute to tonal knowledge acquisition early in development and through late childhood through everyday music listening (Krumhansl & Keil, 1982; Trainor & Trehub, 1992; Trainor & Trehub, 1994). An important challenge for future research will be to determine whether developing infants and children infer tonal prominence through frequency of occurrence and metrical strength as adults did in the present experiments. A related and crucial challenge to support a statistical learning account of tonal knowledge acquisition will be to determine whether exposure to statistical regularities in music results in the formation of long-term tonal schemas. Infants and children would make ideal participants for this purpose, as young listeners have been demonstrated to readily acquire and maintain new musical knowledge even when adults do not (Hannon & Trehub, 2005b). Thus, exposing young listeners to training stimuli with certain statistical properties might cause significant and long lasting changes to their tonal schemas.

Long-term learning was unlikely to have occurred in the present experiments due to the use of the whole tone scale. In the whole tone scale no two pitches are related by the simple ratios 3:2 or 4:3, both of which are high on the stability hierarchy of Western Keys. In neural networks, simple ratio relationships are easier to learn because simple ratios form natural points of stability in the auditory system (Large, 2010). As a result, scales that don’t contain simple ratios, such as the whole tone scale, are less likely to result in long-term learning. The likely facilitation of tonal learning for scales containing simple ratios indicates that while experience is important in musical knowledge acquisition, acquisition is biologically constrained. Previous authors have tended to take one side of music's nature-nurture debate, advocating either for biology (Hemholtz, 1863) or experience (Krumhansl, 1990). Given the evidence in support of both, it is probably more appropriate to view musical knowledge as stemming from an interaction between biological predispositions and musical experience. Accordingly, future research on long-term tonal learning should utilize scales that are rare in Western music and contain simple ratios such as the Bohlen-Pierce scale used in Loui et al. (2010).

Finally, the present findings are inconsistent with the perceptual independence of pitch and temporal structure, as previously advocated by Palmer and Krumhansl (1987). By their definition, Palmer and Krumhansl (1987) observed perceptual independence of pitch and temporal structure because combinations of the two dimensions simply added together to produce participants' melody completeness judgments. Our method of counterbalancing between Familiarizations A and B and our use of the unfamiliar whole tone scale eliminates the possibility that participants' ratings were influenced by previous experience. Nevertheless, metrically strong pitches were given higher ratings than weak pitches. This shows that even though pitch and temporal structure are independent in listeners' perception of melody completion (Palmer & Krumhansl, 1987), meter is capable of influencing the perception of pitch stability. This reinforces the notion that pitch and temporal structure interact at some stage of processing (Peretz & Kolinsky, 1993).

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

Social/Behavioral IRB – Expedited Review

Continuing Review Approved

NOTICE TO ALL RESEARCHERS:

Please be aware that a protocol violation (e.g., failure to submit a modification for any change) of an IRB approved protocol may result in mandatory remedial education, additional audits, re-consenting subjects, researcher probation suspension of any research protocol at issue, suspension of additional existing research protocols, invalidation of all research conducted under the research protocol at issue, and further appropriate consequences as determined by the IRB and the Institutional Officer.

DATE: November 15, 2010

TO: **Dr. Erin Hannon**, Psychology

FROM: Office of Research Integrity – Human Subjects

RE: Notification of IRB Action by Dr. Charles Rasmussen, Co-Chair

Protocol Title: Adult's Perception of Temporal Structure in Music and Speech

Protocol #: 0710-2498

Continuing review of the protocol named above has been reviewed and approved.

This IRB action will reset your expiration date for this protocol. The protocol is approved for a period of one year from the date of IRB approval. The new expiration date for this protocol is November 14, 2011
.

PLEASE NOTE:

Attached to this approval notice is the **official Informed Consent/Assent (IC/A) Form** for this study. The IC/A contains an official approval stamp. Only copies of this official IC/A form may be used when obtaining consent. Please keep the original for your records.

Should there be *any* change to the protocol, it will be necessary to submit a **Modification Form** through ORI Human Subjects. No changes may be made to the existing protocol until modifications have been approved by the IRB.

Should the use of human subjects described in this protocol continue beyond November 14, 2011, it would be necessary to submit a **Continuing Review Request Form** *60 days* before the expiration date.

If you have questions or require any assistance, please contact the Office of Research Integrity - Human Subjects at IRB@unlv.edu or call 895-2794.

BIBLIOGRAPHY

Bernstein, L. (1973). *The unanswered question.* Cambridge, MA: Harvard University Press.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we ‘‘experienced listeners’’? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*, 100-130.

Burns, E.M.: Intervals, scales, and tuning. In: Deustch, D. (ed.) The Psychology of Music, pp. 215–264. Academic Press, San Diego (1999)

Chandrasekaran, B., Homickel, J., Skoe, E., Nicol, T., & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron,* 64*,* 311-319.

Cohen JD, MacWhinney B, Flatt M, and Provost J. (1993). PsyScope: a new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, and Computers*, 25: 257–271, 1993.

Corriveau, K.H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex,* 45, 119-139.

Creel, S., Newport, E. (2002). Tonal profiles of artificial scales: implications for music learning. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition*, Sydney, 2002. Adelaide: Casual Productions.

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.

Cutler, A. & Butterfield, S. (1992). Rhythmic cues to speech segmentation: evidence from juncture misperception. *Journal of Memory and Language*, 31, 218-236.

Drake, C., Jones, M.R., Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition.* 77, 251-288.

Endress, A. D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica,* 135, 182-190.

Ellis, R.J. & Jones, M.R. The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35(1), 264-280*

Gomez, R.L. (2002). Variability and detection of invariant structure. *Psychological Science*. 13(5), 431-436.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences.* 15(1), 3-10.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology, 50*, 354-377.

Hannon, E.E., & Trehub, S.E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science,* 16 (1), 48-55.

Hannon, E.E., Snyder, J.S., Eerola, T., Krumhansl, C.L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance.* 30(5), 956-974.

Hannon, E.E., & Trehub, S.E. (2005b). Tuning in to musical rhythms: infants learn more readily than adults. *Proceedings of the National Academy of Sciences,* 102 (35), 12639-12643.

Helmholtz, H.L.F. (1863). *On the sensations of tone as a physiological basis for the theory of music*. New York: Dover Publications.

Höhle, B. (2009). Bootstrapping mechanisms in first language acquisition. *Linguistics,* 47(2), 359-382.

Huron, D. (2006). *Sweet anticipation: music and the psychology of expectation.* Cambridge, MA: MIT Press.

Huron, D., Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. *Music Perception.* 13(4), 489-516.

James, J. (1993). *The music of the spheres*. New York: Grove Press.

Järvinen, T. (1995). Tonal hierarchies in jazz improvisation. *Music Perception, 12*(4), 415-437.

Jonaitis, E.M., & Saffran, J.R. (2009). Learning harmony: the role of serial statistics. *Cognitive Science.* 33, 951-968.

Jusczyk, P.W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences.* 3(9), 323-328.

Knopoff, L., & Hutchinson, W. (1983). Entropy as a measure of style: The influence of sample length. *Journal of Music Theory,* 27, 75-97.

Krumhansl, C.L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* 11, 346-374.

Krumhansl, C.L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, 126(1), 159-179.

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory & Cognition*, 10(3), 243-251.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*(4), 334-368.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Large, E.W. (2010). A dynamical systems approach to musical tonality. R. Huys and V. K. Jirsa (Eds.): *Nonlinear Dynamics in Human Behavior* (pp. 193–211). Berlin: Springer-Verlag.

Leman, M. (2000). An auditory model of the role of short-term memory in probe tone ratings. *Music Perception*, *17*(4), 481-509.

Leong, V., Hämäläinen, J., Soltész, F., & Goswami, U. (2011). Rise time perception and detection of syllable stress in adults with developmental dyslexia. *Journal of Memory and Language.* 64, 59-73.

Lerdahl, F., & Jackendoff, R. (1996). *A generative theory of tonal music*. Cambridge: MIT Press.

Loui, P., Wessel, D.L., & Hudson Kam, C.L. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception,* 27(5), 377-388.

Lynch, M.P., Eilers, R.E., Oller, K., & Urbano, R.C. (1990). Innateness, experience, and music perception. *Psychological Science*, 1(4), 272-276.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, *20*, 1035-1041.

Meyer, L. B. (1973). *Explaining music: Essays and explorations*. Chicago: University of Chicago Press.

Morey, R.D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*. 4(2). 61-64.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phases. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 116-126.

Patel, A. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674-681.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *The Quarterly Journal of Experimental Psychology*, 46 (2), 301-325.

Povel, D., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2(4), 411-440.

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1598-1617.

Skoe, E. & Kraus, N. Hearing it again: on-line subcortical plasticity in humans. *Plos One*. 5(10), 1-9.

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.

Schellenberg, E.G., & Trehub, S.E. (1996). Natural musical intervals: evidence from infant listeners. *Psychological Science.* 7(5), 272-277.

Trainor, L. T., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 394-402.

Trainor, L. T., & Trehub, S. E. (1994). Key membership and implied harmony in western tonal music: Developmental perspectives. *Perception & Psychophysics, 56*(2), 125-132.

Trehub, S.E., Schellenberg, G., & Damenetsky, S.B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 965-975.

Trehub, S. E., & Thorpe, L. A. (1989). Infants perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology, 43*(2), 217-229.

Winkler, I., Haden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences, 106*(7), 2468-2471.

Youngblood, J. E. (1958). Style as information. *Journal of Music Theory, 24*, 24-35.

VITA

Graduate College
University of Nevada, Las Vegas

Matthew Rosenthal

Degrees:
Bachelor of Science, Psychology, 2008
Indiana University, Bloomington

Master of Arts, Psychology, 2012
University of Nevada, Las Vegas

Special Honors and Awards:
Dynamic Field Theory workshop travel grant
Society for Music Perception and Cognition travel grant

Thesis Title: The role of metrical structure in tonal knowledge acquisition

Thesis Examination Committee:
Chairperson, Erin Hannon, Ph. D.
Committee Member, Joel Snyder, Ph. D.
Committee Member, Mark Ashcraft, Ph. D.
Graduate Faculty Representative, Gabriele Wulf, Ph. D.