

# **Unintentional Learning of Musical Pitch Hierarchy**

By

Anja-Xiaoxing Cui

A thesis submitted to the Graduate Program in Psychology

in conformity with the requirements for the

Degree of Master of Science

Queen's University

Kingston, Ontario, Canada

August, 2014

Copyright ©Anja-Xiaoxing Cui, 2014

## Abstract

It has been proposed that our knowledge about music is acquired passively, without intention, through exposure to music (Zatorre & Salimpoor, 2013). Research has shown that the mental tonal hierarchy of a listener, as assessed by the probe tone technique, often reflects the pitch hierarchy of the music to which the listener has been exposed (Castellano, Bharucha, & Krumhansl, 1984; Lantz, Kim, & Cuddy, 2013). The pitch hierarchy is reflected by the frequency with which each pitch class occurs, i.e., a first-order probability system. However, research concerned with the acquisition of statistical rules in music has only explored learning of second-order probability systems (Loui, Wessel, & Hudson Kam, 2010).

In a series of experiments, I aimed to explore the possibility of unintentional learning of musical pitch hierarchy, a first-order probability system. For this purpose, I assessed participants' sensitivity to frequency information contained in brief excerpts of a novel musical system. I also assessed whether short exposure to a novel musical system would influence participants' representation of pitch hierarchy, and whether participants would be able to distinguish the novel musical system from another musical system to which they were not exposed.

Participants with more music training exhibited higher sensitivity to frequency information contained in a novel musical system than participants with little music training. The sensitivity to frequency information was correlated with the amount of music training that participants had received. However, the mental representation of pitch hierarchy of participants with little music training was more influenced by the pitch hierarchy of the musical system to which they were exposed, whereas the mental representation of pitch hierarchy participants with more music training remained unchanged.

The results suggest that unintentional learning of musical pitch hierarchy is influenced by the amount of music training participants have received. Overall, these findings provide evidence supporting the third proposition of the Theory of Tonal Hierarchies in Music (Krumhansl & Cuddy, 2010), which states that listeners are able to quickly adapt to the tonal hierarchies of unfamiliar music.

## Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Lola L. Cuddy, who replied many months ago to an email from an undergraduate student from Germany. Thank you for giving me the opportunity to work with you: I have enjoyed it tremendously, and I look forward to working with you on my next projects!

I would like to thank my co-supervisor, Dr. Niko F. Troje, for his advice on myriad matters, as well as Dr. Mary C. Olmstead for her insightful comments, and Dr. Jordan L. Poppenk for stepping in at such short notice.

I would also like to thank Dr. John R. Kirby and Dr. Meredith L. Chivers for their time and contribution to the defense of this thesis. I am indebted to Frank Jiang and Kristen Silveira of the Music Cognition Lab, who have contributed their time to helping me run my experiments.

I am grateful to my friends back home for enduring my numerous complaints about Canadian weather, and my friends here who have become my second family.

Finally, I would like to thank Christina Guthier and Yoshiaki Ko for being my remote lab mates, roommates, and cheer squad, my parents Yue Yao and Dr. Xudong Cui for saying that you're not worried but then asking if I eat well, and my sister Susanne-Xiaochen Cui for being my best friend: No one in the world would know from reading "daaaaa dadadaaa dadadaaa dadadaaa daaaaa daa daa daa daa" that Tchaikovsky's violin concerto is playing in my head but you.

## Table of Contents

Abstract.....	ii
Acknowledgements.....	iv
Table of Contents.....	v
List of Tables.....	vii
List of Figures.....	viii
Glossary.....	ix
Chapter 1: Introduction.....	1
The Probe Tone Technique.....	5
A Novel Musical System.....	7
Music Training.....	9
My Experiments.....	11
Chapter 2: Experiments.....	14
Experiment I.....	14
Methods.....	14
Participants.....	14
Stimuli.....	15
Procedure.....	16
Results.....	17
Discussion.....	20
Experiment II a & II b.....	21
Methods.....	21

Participants.....	21
Participants in Experiment II a .....	21
Participants in Experiment II a .....	22
Stimuli.....	22
Stimuli for probe tone ratings .....	23
Stimuli for exposure phase.....	24
Stimuli for two-alternative forced-choice task .....	25
Procedure .....	27
Data Analysis .....	28
Results.....	30
Mixed-model ANOVA .....	30
Exposed/not exposed .....	30
Match/mismatch.....	33
Classification.....	34
Discussion.....	36
Chapter 3: General Discussion.....	39
References.....	47
Appendix I: Music Training Related Descriptors of Sample.....	52
Appendix II: Descriptors of Sample Not Related to Music Training .....	54
Appendix III: Calculation of Dependent Variables and Average Probe Tone Profiles .....	56
Appendix IV: Research Ethics Approval.....	58

## **List of Tables**

Table 1. Overview of my experiments.....	13
Table 2. Pre-composed stimuli pool .....	16

## List of Figures

Figure 1. The probe tone technique .....	6
Figure 2. Relative frequency distribution for Hypophrygian and Lydian mode (Huron and Veltman, 2006).....	9
Figure 3. Probe tone rating.....	17
Figure 4. Average probe tone profile for musicians and nonmusicians for the third phrase of Benedictu es, Domine .....	19
Figure 5. The two stimuli selected as probe tone contexts for experiment II a and b .....	23
Figure 6. Frequency distribution of pitch classes for the Hypophrygian (A) and Lydian mode (B) displayed as probability of occurrence.....	26
Figure 7. Classification task.....	28
Figure 8. Analysis of Experiment II a and Experiment II b.....	30
Figure 9. Results from exposed/not exposed analysis .....	32
Figure 10. Results from match/mismatch analysis .....	35



## Glossary

*Chromatic scale:* A musical scale with 12 pitches, the frequencies of which are based on 12 logarithmically even divisions of an octave (Loui, Wessel, & Hudson Kam, 2010).

*Diatonic scale:* A major or minor scale. Western art music has been based on diatonic scales for most of the past 300 years (Encyclopædia Britannica, 2014; Krumhansl, 1990).

*Event hierarchy:* The rank order of events. The rank is determined by the frequency of occurrence of each event, such that the event occurring the most is placed at the top of the hierarchy, and the event occurring the least is placed at the bottom of the hierarchy.

*First-order probability:* The probability with which an event occurs, independent of previous events (Miller & Selfridge, 1950). E.g. the probability that event B occurs, given that event A just occurred.

*Frequency distribution, or event frequency distribution:* The frequency of occurrence for each pitch ordered along the chromatic scale. The frequency distribution determines the event hierarchy such that the event occurring with the highest frequency is at the top of the hierarchy, and the event occurring with the lowest frequency is at the bottom of the hierarchy.

*Frequency of occurrence:* The number of times with which an event occurs. In this thesis, the number of times with which a certain pitch occurs.

*Generated stimulus:* A stimulus generated based on a pre-defined probability profile.

*Major keys:* All 12 chromatic transpositions of the major scale, with the most important pitch – the tonic – being C, if the seven pitches C, D, E, F, G, A and B are used (Encyclopædia Britannica, 2014). This is reflected by the tonal hierarchy, which places the tonic at the top of the hierarchy (Krumhansl, 1990).

*Minor keys:* All 12 chromatic transpositions of the minor scale, with the most important pitch – the tonic – being A, if the seven pitches C, D, E, F, G, A and B are used (Encyclopædia Britannica, 2014). This is reflected by the tonal hierarchy, which places the tonic at the top of the hierarchy (Krumhansl, 1990).

*Mode:* A type of a scale. Each mode can be characterized by the probability profile describing the frequency of occurrence of each pitch.

*Pre-composed stimulus:* Stimulus based on pre-existing chants.

*Probability profile:* The relative frequency of occurrence of each pitch ordered along the chromatic scale used as probability of occurrence to generate stimuli.

*Probe tone method:* Method developed by Krumhansl and Shepard (1979) to assess the mental hierarchy of pitches.

*Second-order probability:* Probability with which an event occurs dependent on the previous event, i.e., the probability with which event B occurs after event A (Miller & Selfridge, 1950).

*Tonal hierarchy:* The hierarchy of pitches in a musical system. The tonal hierarchy determines the rank order of relative importance of each pitch. The tonal hierarchy of diatonic scales (major or minor scales) assigns the highest importance to the tonic, i.e., the root (or octave), and the lowest to nondiatonic pitches, as described in music theory. It correlates highly with the event hierarchy of corpora of works composed in diatonic scales. The tonal hierarchy is represented mentally in listeners, and is thus a psychological fact (Krumhansl & Cuddy, 2010).

## Chapter 1: Introduction

How do we acquire the implicit musical knowledge that we use to make intuitive judgments about music? How do we “know”, from listening alone, that one piece is written in a major key and another in a minor key?

Major and minor (the two dominant modes in the Western music idiom) can both be described in terms of the first-order probabilities with which pitch classes occur (Krumhansl, 1990). The first-order probabilities form the relative event frequency distribution which describes the number of occurrences for each pitch class. Researchers have proposed that the abstraction of those probabilities through unintentional, i.e., passive, statistical learning is involved in forming our knowledge about the music around us (Zatorre & Salimpoor, 2013).

In the context of music and musical knowledge, i.e., the internal, mental representation of distributional qualities of music, a mechanism that serves to retrieve statistical information could help to build an internal representation of the conventional structures of music. The relative importance of a pitch class within a musical context is reflected in the frequency of occurrence of that pitch class in pieces of the same context (Krumhansl, 1987). For instance, in a C major piece, the pitch class “C” is assumed to be the most important pitch, and it occurs often. In contrast, the pitch class “F #” is then assumed to be relatively unimportant, and it occurs less.

The tonal hierarchy determines the rank order of pitch classes within a given musical system, such as major or minor. The rank of a pitch class is determined by its relative importance, and is reflected by its frequency of occurrence. The internal representation of this hierarchy has been proposed to reflect the listener’s long-term exposure to a musical system (Bharucha, 1984; Deutsch, 1984). Krumhansl (1990) suggested that musical knowledge might therefore be gained through abstraction of the statistical regularities found in music (statistical

learning). According to this theory, our sense of major and minor is gained through exposure to music written in major and minor. We listen to music, abstract regularities found in it, and form an internal representation of this type of music. Because we are exposed to music regularly, and abstraction of the regularities is supposed to take place unintentionally, it is assumed that each person has an internal representation of the tonal hierarchy of the music that he or she was exposed to.

Krumhansl and Cuddy (2010) formulated this as the first proposition of their Theory of Tonal Hierarchies in Music: Tonal hierarchies are psychological facts, i.e., each person has an internal representation of tonal hierarchy.

It should be noted that different styles of music most likely have different tonal hierarchies. This means that the tonal hierarchy cannot be explained solely by physical properties of sound (for an elaboration on how different models of tonality based on physical properties compare to tonal hierarchy of major and minor keys, see Krumhansl, 1990, pp. 55 – 62). The harmonic series (the sequence of multiples of the base frequency) for instance would place more importance on the minor seventh (7 times the base frequency) than the major second (9 times the base frequency) or perfect fourth (11 times the base frequency). However, the tonal hierarchy of major keys as described by Krumhansl and Cuddy (2010) places more importance on the perfect fourth than on the major second or on the minor seventh.

There is a train of thought in psycholinguistic research regarding the connection between abstraction of statistical regularities in speech and learning of language similar to the one in music psychology research regarding the connection between abstraction of statistical regularities in music and gaining of musical knowledge. Some psycholinguists propose that the learning of language involves the abstraction of statistical regularities found in speech we hear

(Saffran, Aslin, & Newport, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). They argue that linguistic knowledge might be gained through abstraction of the statistical regularities found in language: We listen to language, abstract regularities found in it, and form an internal representation of the language.

In fact, previous research has found that statistical information can be retrieved from an array of stimuli types. These include the visual domain (Fiser & Aslin, 2002) and complex non-verbal auditory stimuli (Tillman & McAdams, 2004). Statistical learning has also been shown in cotton-top tamarins, a non-human primate (Hauser, Newport, & Aslin, 2001). This highlights the multimodality and potential of passive statistical learning to be a general cognitive capacity of the human brain (see Thiessen & Erickson, 2013).

Evidence supporting the view that exposure influences musical knowledge stems from developmental studies (Trainor & Trehub, 1994), studies comparing musicians with non-musicians (Oram & Cuddy, 1995), and cross-cultural studies (Castellano, Bharucha, & Krumhansl, 1984; Kessler, Hansen, & Shepard, 1984; Lantz, Kim, & Cuddy, 2013). However, exposure to music in those studies was not directly manipulated. Differences in exposure were assumed based on participants' age (it was assumed that older people have had more exposure), profession (it was assumed that musicians have had more exposure), and cultural background (it was assumed that there was exposure to the music of the culture in which the participant grew up in, but no exposure to music of another culture).

The question that presents itself here is whether exposure to a musical system for a short amount of time, within a single experiment, will be sufficient to establish knowledge about this musical system. This has been suggested by Krumhansl and Cuddy (2010) in the third and last

proposition of their Theory of Tonal Hierarchies in Music, which states that “listeners rapidly adapt to style-appropriate tonal hierarchies even if the style is unfamiliar” (p. 80).

A recent study by Loui, Wessel, and Hudson Kam (2010) demonstrated passive statistical learning of second-order probabilities (transitional probabilities), reminiscent of linguistic studies (Saffran et al., 1996; Saffran et al., 1997). With my experiments, I would like to expand our knowledge of passive statistical learning of music, by extending research to passive statistical learning of a set of first-order probabilities, such as pitch hierarchy, i.e., the frequency distribution of pitch classes.

The appeal of first-order probabilities lies in the parsimony of a model that relies on them. I do not expect such a model to fully describe our musical experience, but I trust it to be a good starting point. Furthermore, in work by music theorists, first-order probabilities have been proposed as a basis for key finding algorithms (Huron & Veltman, 2006; Krumhansl, 1990; Temperley, 2007).

The objective of this series of experiments was to investigate whether unintentional learning of first-order probabilities (such as pitch hierarchy) occurs after short exposure, by having participants listen to a novel musical system (that simulates exposure to music in our everyday life), and asking whether this exposure influenced the internal representation of the musical structures afterwards. I also investigated whether music training plays a role in shaping passive statistical learning of novel musical material.

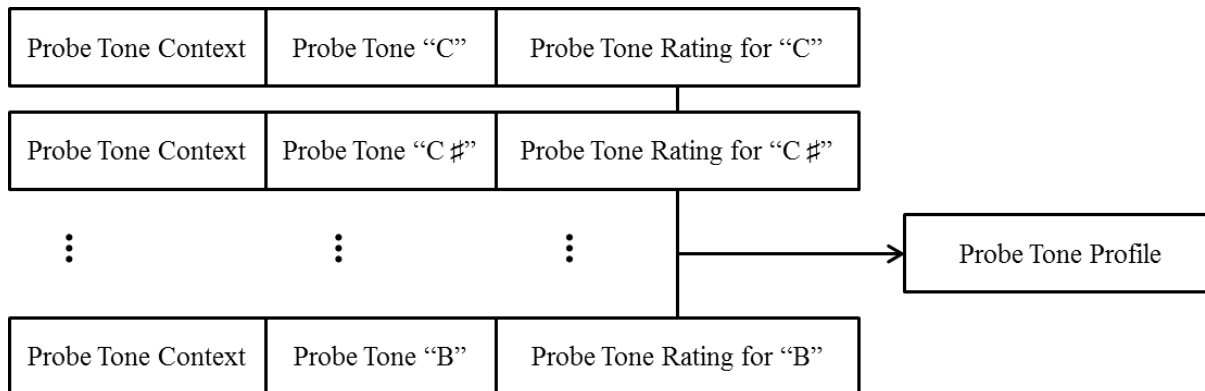
With this research, I want to bridge the gap that currently exists between research on statistical learning in the musical domain and the theorists who posit that our musical knowledge is gained through statistical learning. These theorists refer to research with first-order probability systems (see Castellano et al., 1984; Kessler et al., 1984; Lantz et al., 2013). The referenced

studies examined the representation of tonal hierarchy in listeners using nondiatonic stimuli. The representation of tonal hierarchy as assessed by the probe tone technique (explained in more detail in the following section) was compared to the pitch hierarchy, which is a first-order probability system. The research on statistical learning in the musical domain on the other hand has only explored learning of second-order probability systems, i.e., systems defined by transitions of pitches or chords (Rohrmeier & Rebuschat, 2012). Studies using second-order probability systems borrowed methodology from linguistic studies, which were mainly concerned with systems defined by transitional rules (Saffran et al., 1996).

In the remainder of this chapter, I will describe the experimental paradigm I used to assess the representation of tonal hierarchy of my participants, and the novel musical system they were exposed to. I also elaborate on differences between participants with more music training and participants with less music training. This chapter ends with a brief overview of my experiments.

### **The Probe Tone Technique**

Studies investigating musical knowledge have made use of the probe tone technique developed by Krumhansl and Shepard (1979). In this paradigm, the listener is asked to evaluate the musical goodness-of-fit of a probe tone to a musical context (the probe tone context). The probe tone context can be established by tone sequences, chords, or cadences (Cuddy & Badertscher, 1987; Krumhansl & Kessler, 1982; Oram & Cuddy, 1995). The probe tone context defines the musical system, for which the internal representation of the structure of this musical system, i.e., knowledge about this musical system, is then activated. The set of probe tone ratings ordered chromatically along the scale for all 12 chromatic pitch classes (from C to B) is called a probe tone profile. The probe tone technique is visualized in Figure 1.



*Figure 1.* The probe tone technique. The probe tone ratings for the twelve probe tones representing the twelve pitch classes of the chromatic scale to the same probe tone context form the probe tone profile for that context. Each row represents a trial. It should be noted that the probe tone ratings are usually not collected for probe tones in chromatic but rather in random order.

The listener is asked to rate how well the probe tone fits with the probe tone context (goodness-of-fit) on a numerical scale. The probe tone profile is assumed to be a quantification of the listener’s representation of tonal hierarchy. In my experiments, the probe tone contexts are monophonic melodies. Participants are asked to assess the musical goodness-of-fit of probe tones on a scale from 1 (“doesn’t fit at all”) to 7 (“fits very well”).

In the case of the Western tonal-harmonic idiom, which consists of the major and minor keys, the average probe tone profile has been found to be similar to the frequency distribution of pitch classes found in pieces or corpora of pieces, i.e., the event hierarchy in these corpora, composed in this idiom (Krumhansl, 1985; Krumhansl, 1990; Temperley, 2010). Krumhansl (1985) reported that probe tone profiles (the representation of tonal hierarchy in listeners) correlated highly with the pitch class count performed by Knopoff and Hutchinson (1983). Knopoff and Hutchinson (1983) counted the frequency of occurrence of pitch classes in vocal melodic lines from compositions by Schubert, Mozart, Hasse, and Strauss. Thus, the event hierarchy of these pieces can be regarded as approximation of the tonal hierarchy as an objective property of music.



I ran correlation analyses with the tabulation by Knopoff and Hutchinson (1983) and the standardized key profile (formed by averaging transposed probe tone profiles obtained for contexts of different keys, as described in Krumhansl & Cuddy, 2010) for pieces written in major keys. As both the pitch class count and the standardized key profiles list values for each of the twelve chromatic pitches, the sample size was  $n = 12$  for each analysis. The correlation coefficients from these analyses ranged between  $r = .84$  (Mozart arias and songs) to  $r = .93$  (Strauss pieces) (Schubert lieder:  $r = .88$ , and Hasse pieces:  $r = .88$ , all  $ps < .05$ ).

To put it another way, the tonal hierarchy of a musical system as perceived by listeners (the internal representation of pitch class hierarchy as assessed by the probe tone technique) is closely related to the event hierarchy (the frequency distribution within a piece; see Bharucha, 1984) of a piece composed in this musical system. A musical system can thus, to some extent, be defined by the frequencies with which pitch classes occur in a musical stimulus, i.e., the first-order probabilities.

In the Theory of Tonal Hierarchies in Music (Krumhansl & Cuddy, 2010), this is reflected in the second proposition: Tonal hierarchies are musical facts also, i.e., they are related to objective properties of music. Based on this proposition, the high correlations between the standardized key profile and the pitch class count (Krumhansl, 1985) show that the probe tone technique is suited to quantify the mental representation of tonal hierarchy.

### **A Novel Musical System**

The musical systems used in my experiments, the Hypophrygian and Lydian mode, are unfamiliar to the general Western audience; they would likely only be familiar to scholars who specialize in medieval church music and Gregorian chants. The general Western audience can be expected to have had exposure to pieces written in major and minor keys, as these have been

used during the past 300 years (Encyclopædia Britannica, 2014). An analysis of a random sample of over 100 Billboard 100 songs from 1958 to 1991 by Burgoyne, Wild, and Fujinaga (2013) revealed predominant usage of diatonic systems. The choice of nondiatonic systems thus ensures that participants have had little to no exposure to the stimuli. This removes confounds found in studies using excerpts from more popular classical pieces (Schellenberg, Peretz, & Vieillard, 2008; Szpunar, Schellenberg, & Pliner, 2004).

Both the Hypophrygian and Lydian system use the same pitch classes as the diatonic scales. They differ from each other, and from the major and minor keys, by the frequencies with which each of these pitch classes occur. This means that their pitch hierarchies are different.

At the same time, the individual pitches comprising the Hypophrygian and Lydian mode are familiar to the general Western audience, as the major and minor keys use them as well. This characteristic makes them more ecologically valid than the Bohlen-Pierce scale used in experiments similar to mine; the Bohlen-Pierce scale uses different pitches based on 13 divisions of a tritave in contrast to the 12 divisions of an octave used in major and minor keys (Loui & Wessel, 2008; Loui et al., 2010; Loui & Schlaug, 2012).

The frequency distributions for the Hypophrygian and Lydian mode were described by Huron and Veltman (2006) based on a sample of 98 chants from the *Liber usualis* (Benedictines of Solesmes, 1961), which contains over 2000 medieval chants. The frequency distribution of the sample was found to be useful for modal classification of other chants of the *Liber usualis* (Huron & Veltman, 2006). Using the frequency distribution as basis of an algorithm, Huron and Veltman (2006) were able to assign the same mode to chants that the monastic scholars gave who collected the chants in the *Liber usualis*.

Figure 2 depicts the event frequency distribution of the two modes as described by Huron and Veltman (2006). If one were to describe each of the modes by the pitch classes that occur most often, Hypophrygian could be described as often using “F” and “G”, and Lydian could be described as often using “C” and “A”.

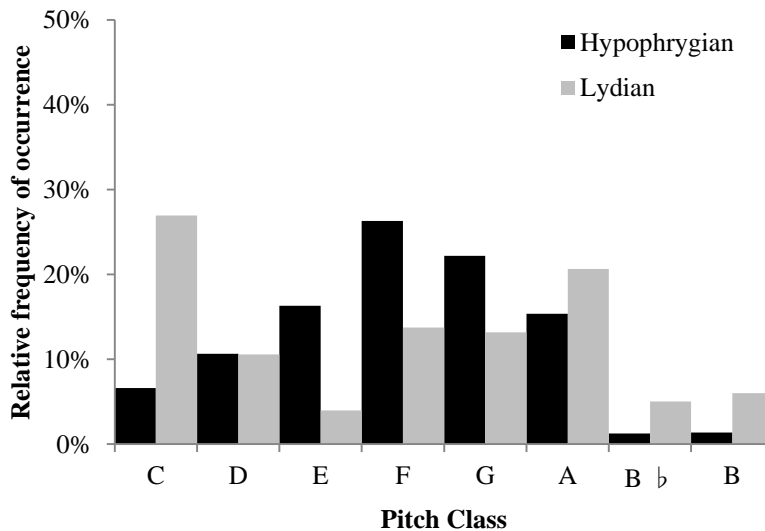


Figure 2. Relative frequency distribution for Hypophrygian and Lydian mode (Huron & Veltman, 2006). Note the distinct “peaks” in the distribution at different pitch classes for the two modes.

### Music Training

As previously mentioned, some studies have attempted to model exposure to music by testing participants with different levels of music training. For example, in a study using the probe tone technique by Oram and Cuddy (1995), participants who had extensive music training (Grade IX Royal Conservatory of Music) gave probe tone ratings that reflected greater influence by the frequency of occurrence compared to participants with no formal music training.

Furthermore, probe tone ratings by participants with more training tended to reflect the influence of tonal hierarchy of major or minor keys more than did probe tone ratings by participants with no formal music training.

This pattern is similar to the one found by Krumhansl and Shepard (1987). Participants in their study tended to respond in the same direction, but probe tone profiles by participants with more music training were more distinct than probe tone profiles by participants with little or no music training. For the sake of convenience, I will refer to participants with more or extensive music training as musicians, and to participants with less, little or no music training as nonmusicians.

A separate line of research has established neurocognitive differences between musicians and nonmusicians. A number of studies have used event related potential (ERP) measures to assess differences in brain waves. In studies with adults, musicians demonstrated enhanced early right anterior negativities (ERAN), which reflect departures from musical regularities established internally in long-term format (Koelsch, Schmidt, & Kansok, 2002; Koelsch, 2009; Koelsch, 2013). Children with music training also had enhanced ERAN compared to children without music training (Jentschke & Koelsch, 2009).

Further, a recent semi-longitudinal study found that music training modified the mismatch negativity response (MMN) in children (Putkinen, Tervaniemi, Saarikivi, de Vent, & Houtilainen, 2014). The MMN reflects departures from musical regularities established internally in short-term format (Koelsch, 2013). Thus, auditory discrimination may be enhanced by music training (Kraus & Chandrasekaran, 2010).

Both musicians and nonmusicians participated in my experiments, thus adding an additional factor to consider during statistical analyses. Based on the existing research, it can be said that differences are likely to exist, though in what form has yet to be determined. On the one hand, enhanced auditory discrimination in musicians might lead to more distinct probe tone profiles. Enhanced auditory discrimination could also enable musicians to abstract frequency

information more quickly. On the other hand, the extensive training musicians received might lead to an internal representation of musical structure that is less likely to be influenced by short exposure to novel music. Either way, it has been argued that musicians possess a more adaptive auditory system (Kraus & Chandrasekaran, 2010), and better working memory (Pallesen et al., 2010).

It should be noted that because musical enculturation is thought to take place passively, i.e., through mere exposure (Corrigall & Trainor, 2010), nonmusicians are expected to also have a representation of tonal hierarchy (first proposition of the Theory of Tonal Hierarchies in Music, see Krumhansl & Cuddy, 2010). However, differences are likely to exist based on the fact that music training leads to earlier enculturation to musical structure. This has been shown in behavioral studies (Trainor, Marie, Gerry, Whiskin, & Unrau, 2012), as well as neuroscientific studies (Koelsch, 2013).

## **My Experiments**

In a series of experiments, I explored the possibility of abstraction of frequency information to be a mechanism involved in gaining musical knowledge. With this in mind, I wanted to examine the general sensitivity to frequency information in tone sequences of a novel musical system (Hypophrygian or Lydian) as quantified by the probe tone method. I also wanted to examine whether exposure to tone sequences embodying one of the distributions described by Huron and Veltman (2006) would change an initial probe tone profile.

More specifically, I wanted to explore the possibility of a change in the probe tone profile that indicates implicit learning by the participant (Rohrmeier & Rebuschat, 2012). I was interested in whether the probe tone context would activate an internal representation of the structure of the novel musical system after exposure; the latter assessed by the probe tone profile.

I was also interested in seeing whether this musical knowledge can be used to make classificatory decisions, and whether the music training the listener has received would influence those results.

In Experiment I, I tested musicians' and nonmusicians' sensitivity to statistical information (frequency of occurrence of pitch classes) contained in excerpts of a novel musical system using the probe tone technique. This experiment aimed to address the third proposition of the Theory of Tonal Hierarchies in Music (Krumhansl & Cuddy, 2010), which states that listeners are able to quickly adapt to the tonal hierarchies of unfamiliar music. Moreover, this experiment was designed to obtain data for stimulus selection for Experiment II a and Experiment II b.

With Experiment II a and Experiment II b, I wanted to again address the third proposition of the Theory of Tonal Hierarchies in Music (Krumhansl & Cuddy, 2010). The question I asked was whether listeners' representation of tonal hierarchy would be changed after short exposure to a novel musical system.

Using the musical systems described earlier allowed me to address two concerns with existing research: First, by using an unfamiliar musical system, between subject variability in exposure is more controlled than in studies using excerpts from classical pieces (Schellenberg et al., 2008; Szpunar et al., 2004). Second, by using a musical system based on historical music assures some degree of musicality in the stimuli, thereby making it more ecological valid than studies using new types of scales (Loui et al., 2010).

In Experiment II b, I also tested whether exposure to a novel musical system would enhance familiarity of melodies from that musical system compared to melodies from another musical system.

Table 1 provides an overview of the experiments and the questions they address.

Table 1

*Overview of my experiments*

<b>Experiment</b>	<b>Questions</b>
I	Assessment of mental hierarchy of new musical system using the probe tone technique, differences between musicians and nonmusicians, obtaining data for stimuli selection for Experiment II a and Experiment II b
II a	Effects of exposure to tone sequences of a novel musical system on probe tone profiles, differences between musicians and nonmusicians
II b	Effects of exposure to tone sequences of a novel musical system on probe tone profiles (replication of Experiment II a), differences between musicians and nonmusicians (replication of Experiment II a), effects of exposure to tone sequences of a novel musical system on classification of new musical material

## Chapter 2: Experiments

### Experiment I

With Experiment I, I used the probe tone technique to assess musicians' and nonmusicians' sensitivity to statistical information (frequency of occurrence of pitch classes) contained in excerpts of a novel musical system. Furthermore, this experiment was designed to obtain data based on which the stimulus selection for Experiment II a and Experiment II b would take place.

#### **Methods.**

#### ***Participants.***

Ten musicians and 10 nonmusicians were recruited from the student body of Queen's University to participate in Experiment I. Participants were classified as musicians if they held Grade X Royal Conservatory of Music (RCM) certificates or equivalent, and/or were taking university level music classes. Participants were classified as nonmusicians if they had less than five years of formal music training (private lessons and/or institutional classes). Students who indicated interest in participating but did not fall into one of the categories were excluded from participation. On average, musicians had significantly more years of training than nonmusicians (musicians:  $M = 11.80$ ,  $SD = 2.30$ , nonmusicians:  $M = 2.80$ ,  $SD = 0.84$ ,  $t(18) = 11.79$ ,  $p < .001$ ). Participants were compensated \$5 for their time. All participants reported normal hearing. Participants' ages ranged from 18 to 26 years ( $M = 21.87$ ,  $SD = 2.42$ ). For more descriptors of the participants, see Appendix.



### *Stimuli.*

Huron and Veltman (2006) used a sample of the works found in the *Liber usualis* to determine the distribution of pitch classes for each mode by tallying the tones occurring in the respective chants. These mode profiles were tested by assessing the Euclidean distance between those aggregate mode profiles and individual pitch class distributions for two test chants per mode and proved to be useful for mode classification.

Thus, using the test chants as stimuli in my experiment will not only ensure that they are considered Hypophrygian or Lydian based on music theory (the classifications made by the monastic scholars), but that they also represent the Hypophrygian or Lydian distribution according to Huron and Veltman (2006). Note that this representation is not perfect: While the mode classification is possible based on the overall distribution, the frequency distributions found within one test chant may differ from the overall distribution.

As chants consist of a series of phrases, I decided to use phrases rather than the whole chant; phrases are shorter in duration and therefore will reveal less frequency information. At the same time, they are still musically valid statements. Phrases were selected such that the number of notes of the phrase ranged between 20 and 40 notes. Using the four chants mentioned above, I compiled a pool of 20 pre-composed stimuli with the number of notes ranging from 22 to 39. Table 2 describes the pre-composed stimuli pool. *Tenuisti manum* and *Illumina oculos meos* are Hypophrygian chants. *Benedictus es, Domine* and *Gloria et honore* are Lydian chants.

I then proceeded to record these stimuli. For this purpose, I used tones in .aiff format provided in the online archive of the University of Iowa Electronic Music Studios (Fritts, 2013) with frequencies ranging from A<sub>3</sub> (220 Hz) to G<sub>5</sub> (784 Hz). These tones were recorded on a Steinway & Sons B model. This ensured that melodies sounded like they were recorded on a

piano. Tones were converted to .wav format and built to melodies using Audacity. All tones in tone sequences were sounded for 220 ms with the next tone beginning 200 ms after onset of the tone before to create a legato effect. This rendered stimuli ranging in length between 4.42 s and 7.82 s.

Table 2  
*Pre-composed stimuli pool*

<b>Chant name</b>	<b>Page in Liber Usualis</b>	<b>Phrases</b>	<b>Notes in each phrase</b>
Tenuisti manum	593-94	7	22, 23, 31, 30, 27, 34, 27
Illumina oculos meos	1000-01	4	34, 38, 39, 26
Benedictu es, Domine	910	5	31, 38, 22, 26, 25
Gloria et honore	1133-34	4	23, 28, 34, 39

***Procedure.***

Each participant gave probe tone ratings for six pre-composed stimuli serving as probe tone context; thus a total of six probe tone profiles (72 probe tone ratings) per participant were collected. The six pre-composed stimuli were chosen quasi-randomly such that each pre-composed stimulus served as context six times. This meant that six probe tone profiles per stimulus (three by musicians and three by nonmusicians) were recorded.

The probe tone technique I adapted for the experiment was designed as follows: For each probe tone rating, participants were asked to evaluate how well the probe tone fitted into the probe tone context (a pre-composed stimulus) on a Likert-scale from 1 to 7. A score of 1 indicated that the probe tone did not fit in the musical context at all; a score of 7 indicated that the probe tone fitted well in the musical context. The probe tones were sounded for 1000 ms after a gap of 2000 ms between the end of the last tone of the probe tone context and the probe tone itself. Probe tone ratings were collected for the twelve pitch classes comprising a chromatic scale

from C<sub>4</sub> (262 Hz) to B<sub>4</sub> (494 Hz). I used the same .aiff files to create the probe tones that I used for recording the pre-composed stimuli.

Figure 3 illustrates an example probe tone rating trial. The twelve ratings collected for the one pre-composed stimulus serving as musical context were ordered chromatically along the scale (i.e., from C<sub>4</sub> to B<sub>4</sub>) to create a probe tone profile (an array of numbers ranging from 1 to 7) for the probe tone context.

The experiment took place in a sound attenuated chamber. The experimental program was written using MATLAB and presented using a Dell Precision T1500 PC. Participants were instructed to adjust the volume to a comfortable level.

pre-composed stimulus	2 s pause	1 s probe-tone	1 s pause	“How well does the single tone fit in the melody played before?”
-----------------------	-----------	----------------	-----------	--

*Figure 3.* Probe tone rating. The musical context is established by playing a pre-composed stimulus, which is followed by a pause of 2 s. The probe tone is sounded for 1 s, and after a pause of 1 s, participants are asked to rate the musical goodness-of-fit.

### **Results.**

For each participant, a score of their general sensitivity was calculated as follows. First, a correlation coefficient between the probe tone profile and the pitch class frequency count was computed. For this purpose, the pitch class frequency count was first ordered along the chromatic scale from C<sub>4</sub> (262 Hz) to B<sub>4</sub> (494 Hz). The correlation coefficient between the probe tone profile and the pitch class frequency count of the probe tone context was then Fisher  $z$ -transformed to normalize the correlation coefficients. I took the mean of each participant’s six Fisher  $z$ -transformed correlation coefficients to obtain a score of this participant’s general sensitivity to the pitch class distribution as quantified by the probe tone method, henceforth called *GS*.

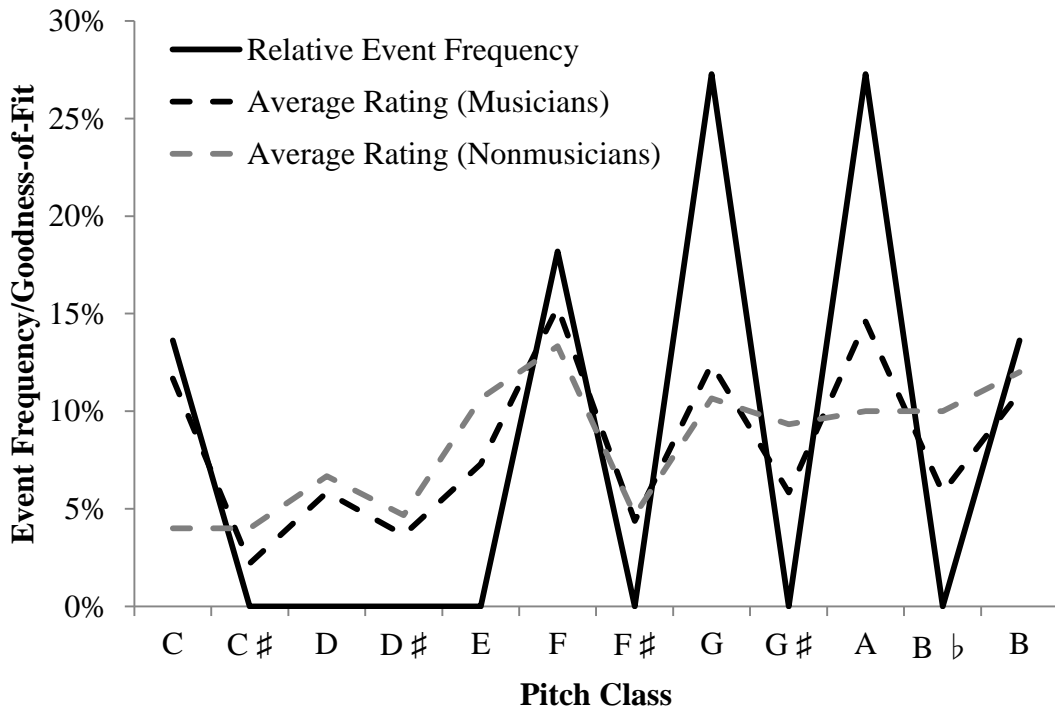
An independent samples  $t$ -test revealed that  $GS$  differed significantly between musicians ( $M = 0.88$ ,  $SD = 0.17$ ) and nonmusicians ( $M = 0.62$ ,  $SD = 0.17$ ), such that musicians had a significantly higher  $GS$  than nonmusicians,  $t(18) = 3.43$ ,  $p = .003$ ,  $d = 1.53$ . To get a better sense of the magnitude of the  $GS$ , the inverse of the Fisher transform was calculated; the Fisher inverse of the  $GS$  can be interpreted as a correlation coefficient. The Fisher inverse of  $GS = 0.88$  is  $GS_{inv} = .71$ , the Fisher inverse of  $GS = 0.62$  is  $GS_{inv} = .55$ . If interpreted as a correlation coefficient, this would indicate a significant correlation for musicians,  $r(10) = .71$ ,  $p = .010$ , and a marginally significant correlation for nonmusicians,  $r(10) = .55$ ,  $p = .064$ . Furthermore, there was a significant correlation of  $r(18) = .66$ ,  $p = .001$ , between  $GS$  and years of music training that participants had received.

A score similar to  $GS$  was also calculated, for which the correlation coefficients between the probe tone ratings and an array of numbers representing the tonal hierarchy (standard probe tone profile for the major keys) as described in Krumhansl and Cuddy (2010) were Fisher  $z$ -transformed and then averaged. This score served as an estimate of participants' tonal assimilation ( $TA$ ). The  $TA$  conveys an idea of how much the tonal hierarchy influenced participants' responses.

An independent samples  $t$ -test revealed that  $TA$  also differed significantly between musicians ( $M = 0.75$ ,  $SD = 0.20$ ) and nonmusicians ( $M = 0.42$ ,  $SD = 0.22$ ), such that musicians had a higher  $TA$  than nonmusicians,  $t(18) = 3.51$ ,  $p = .003$ ,  $d = 1.57$ . To get a better idea about the magnitude of those values, the inverse was again calculated; the Fisher inverse of  $TA = 0.75$  is  $TA_{inv} = .63$ , the Fisher inverse of  $TA = 0.42$  is  $TA_{inv} = .39$ . If interpreted as a correlation coefficient, this would indicate a significant correlation for musicians,  $r(10) = .63$ ,  $p = .028$ , and a nonsignificant correlation for nonmusicians,  $r(10) = .39$ ,  $p = .210$ .

A dependent samples *t*-test revealed that, on average, *GS* ( $M = 0.75$ ,  $SD = 0.21$ ) was significantly higher than *TA* ( $M = 0.58$ ,  $SD = 0.27$ ;  $t(19) = 4.50$ ,  $p < .001$ ,  $d = 0.70$ ).

Figure 4 shows the average probe tone profile for musicians and nonmusicians for the third phrase of *Benedictus, Domine*, expressed as a proportion of the sum of all goodness-of-fit ratings. The average *GS* for musicians to this probe tone context was  $GS = 1.48$ , for nonmusicians  $GS = 0.51$ . The Fisher inverse of  $GS = 1.48$  is  $GS_{inv} = .90$ , the Fisher inverse of  $GS = 0.51$  is  $GS_{inv} = .47$ .



*Figure 4.* Average probe tone profile for musicians and nonmusicians for the third phrase of *Benedictus, Domine*. The average probe tone profile is calculated by averaging the probe tone ratings for each pitch class obtained from musicians or nonmusicians. This figure shows the average probe tone rating by musicians in a dashed black line, and the average probe tone rating by nonmusicians in a dashed grey line. The solid black line depicts the relative event frequency profile of the probe tone context for which the probe tone ratings were obtained (the third phrase of *Benedictus, Domine*).

## **Discussion.**

The significant correlations between the pitch hierarchy (the pitch class frequency count) and probe tone profiles indicated that listeners were sensitive to distribution information that was contained in the short excerpt of music they listened to. Furthermore, there was a significant difference in the general sensitivity between musicians and nonmusicians such that musicians exhibited higher sensitivity, even though musicians have unlikely been previously exposed to the musical style from which the stimuli were drawn.

Musicians' scores revealed that they were more influenced by the major tonal hierarchy than nonmusicians, mirroring results from other studies that showed higher tonal assimilation for musicians (Krumhansl & Shepard, 1987; Oram & Cuddy, 1995). This indicates the adequate use of the probe tone technique. Overall, this experiment showed that the probe tone technique can be used to assess sensitivity to frequency information contained in excerpts of a novel musical system.

The significant sensitivity to frequency information supports the third proposition of the Theory of Tonal Hierarchies in Music (Krumhansl & Cuddy, 2010), which states that listeners are able to quickly adapt to the tonal hierarchies of unfamiliar music. However, the question remains, whether participants are able to abstract the underlying probabilities of pitch class occurrence of multiple stimuli over an extended period of time. Therefore, in Experiment II a and Experiment II b, I assessed participants' mental hierarchy before and after exposing them to multiple tone sequences generated using the same probability profile.

The data obtained in Experiment I were also used to select stimuli for Experiment II a and Experiment II b. I describe the selection process later on.

## **Experiment II a & II b**

With Experiment II a, I aimed to explore whether exposure to a new musical system would influence the mental hierarchy of this system. With Experiment II b, I wanted to examine whether there would be an influence on a mode classification task. Experiment II b had the same procedure as Experiment II a, save for an added part at the end.

In the overlapping part, participants were exposed to stimuli that were generated based on the probability profile of one mode. Before and after exposure, probe tone profiles were obtained using a probe tone context of the same mode, and a probe tone context of a different mode. I regressed the probe tone profiles on the event frequency profiles of the generated stimuli to obtain dependent variables. I analyzed the beta weights from these regressions to investigate changes of influence by exposure.

The additional part of Experiment II b paired stimuli from the two different musical systems. Participants were asked to indicate which stimulus they found more familiar. With this part, I aimed to explore whether participants would differentiate between the mode to which they had been exposed and the mode to which they had not been exposed.

### **Methods.**

#### ***Participants.***

##### *Participants in Experiment II a.*

Ten musicians and 10 nonmusicians participated in Experiment II a. Participants were classified as musicians or nonmusicians according to the same criteria as in Experiment I. None of the participants had participated in Experiment I. Participants were compensated \$8 for their time. All participants reported normal hearing. Participants in this experiment also completed the

Barcelona Music Reward Questionnaire (BMRQ; Mas-Herrero, Marco-Pallares, Lorenzo-Seva, Zatorre, & Rodriguez-Fornells, 2013). The BMRQ was included to provide an estimate of participants' reward experiences with music. Positive reward experiences with music might be concurrent with more engaged listening. On average, musicians had significantly more years of training than nonmusicians (musicians:  $M = 11.20$ ,  $SD = 3.71$ , nonmusicians:  $M = 2.50$ ,  $SD = 1.71$ ,  $t(18) = 7.22$ ,  $p < .001$ ). Participants' age ranged between 18 and 27 years ( $M = 21.61$ ,  $SD = 2.57$ ). For more descriptors of the participants, including the BMRQ, see Appendix.

#### *Participants in Experiment II b.*

Twelve musicians and 10 nonmusicians participated in Experiment II b. Participants were classified as musicians or nonmusicians according to the same criteria as in Experiment I. None of the participants had participated previously in Experiment I or Experiment II a. Participants were compensated \$10 for their time. On average, musicians had significantly more years of training than nonmusicians (musicians:  $M = 10.75$ ,  $SD = 3.41$ , nonmusicians:  $M = 2.78$ ,  $SD = 1.06$ ,  $t(20) = 7.19$ ,  $p < .001$ ). All participants reported normal hearing. Participants in this experiment also completed the BMRQ (Mas-Herrero et al., 2013). Participants' age ranged between 17 and 33 years ( $M = 21.13$ ,  $SD = 4.19$ ). For more descriptors of the participants, including the BMRQ, see Appendix.

#### *Stimuli.*

There were three different classes of stimuli for Experiment II a and Experiment II b. I chose two of the pre-composed stimuli that were used in Experiment I to be used as probe tone contexts in Experiment II a and Experiment II b. Participants also listened to stimuli that were generated based on pre-defined probability profiles (generated stimuli). The third class of stimuli



was only used in Experiment II b. These stimuli were also selected from the pre-composed stimuli that were used in Experiment I.

*Stimuli for probe tone ratings.*

Two of the pre-composed stimuli used in Experiment I were selected for use in Experiment II a and Experiment II b as probe tone contexts. To choose the excerpts, I computed two performance scores per excerpt from the data of Experiment I. The first score was computed by averaging Fisher  $z$ -transformed correlations of pitch class frequency count of the probe tone context with probe tone profiles across participants (averaging subjects, *PSS*). The second score was computed by averaging probe tone profiles before correlating the averaged probe tone profile with the pitch class frequency count of the probe tone context (averaging ratings, *PSR*). Both context performance scores (*PSS* and *PSR*) correlated highly: For the musicians  $r_{PSS \times PSR}(18) = .97, p < .001$ , and for the nonmusicians  $r_{PSS \times PSR}(18) = .93, p < .001$ .

The excerpts for the probe tone ratings in Experiment II a and Experiment II b were chosen such that the performance scores did not differ significantly between the two groups. After excluding the excerpts whose performance scores differed significantly between the groups, one excerpt per mode was chosen that had the least variability. This was determined by the standard deviation of the six correlations between the probe tone profile and the pitch class frequency count, which were collected for each pre-composed stimulus. Based on this selection process, two probe tone contexts were chosen: The first phrase of *Tenuisti manum* and the second phrase of *Gloria et honore* (see Table 2). The phrases are displayed in traditional staff notation in Figure 5.

Hypophrygian: "Tenuisti manum" Phrase 1



Lydian: "Gloria et honore" Phrase 2



*Figure 5.* The two stimuli selected as probe tone contexts for experiment II a and b. The Hypophrygian stimulus was the first phrase from *Tenuisti manum*; the Lydian stimulus was the second phrase from *Gloria et honore*.

*Stimuli for exposure phase.*

A second set of stimuli were generated using MATLAB to follow the pitch class hierarchy of the Hypophrygian and Lydian mode as described by Huron and Veltman (2006). For each mode, the values for the 12 pitch classes – as described by Huron and Veltman (2006) – were raised by an exponent of 2 and expressed as a percentage of the sum of all these values. This manipulation maintains the pitch class hierarchy. However, this hierarchy is now slightly exaggerated (see Smith & Schmuckler, 2004). Both the original pitch class hierarchies and the exaggerated version are shown in Figure 6; note that the rank order of the pitch classes is maintained.

This exaggeration was introduced because of previous work at the Music Cognition Lab, which showed that melodies from exaggerated profiles are more familiar than melodies from less exaggerated versions after exposure (Collett, 2013). The percentages were used as a probability profile, i.e., as probability of occurrence for each pitch class (see Figure 6). The tone sequences were matched in the number of tones, i.e., in duration to the pre-composed stimuli to mimic the length of actual phrases.

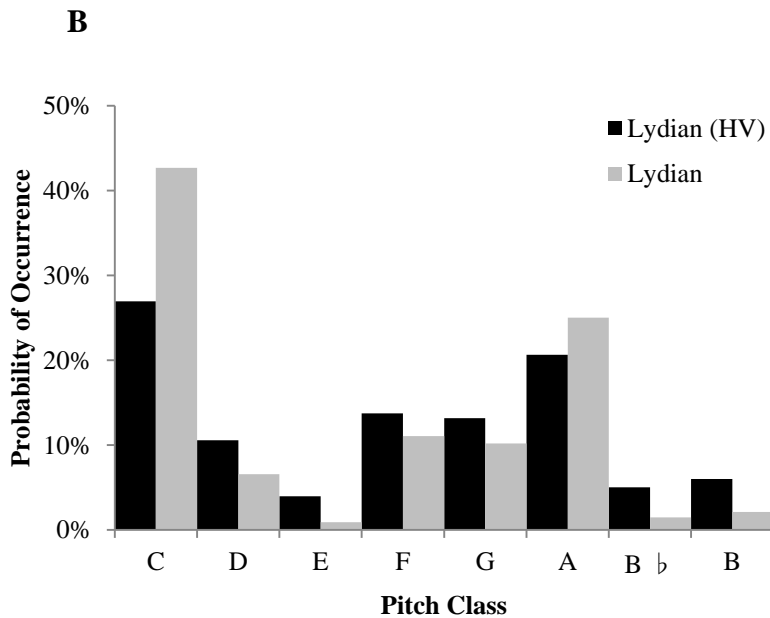
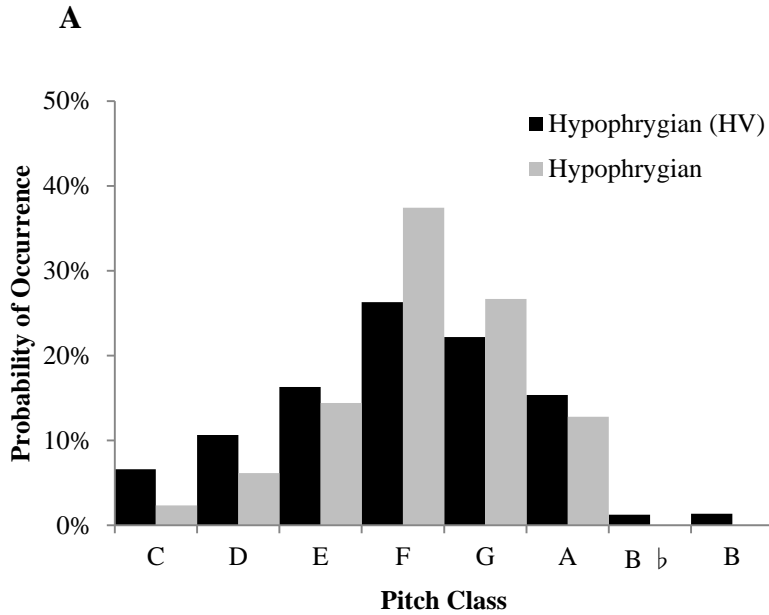
Again, I used tones in .aiff format provided in the online archive of the University of Iowa Electronic Music Studios (Fritts, 2013). The generated stimuli used tones with frequencies

ranging from C<sub>4</sub> to B<sub>4</sub>. These were converted to .wav format using Audacity and built to melodies using MATLAB. All tones in the generated tone sequences were sounded for 220 ms with the next tone beginning 200 ms after onset of the tone before to create a legato-feeling. This was the same procedure used to record the stimuli from the *Liber usualis*.

The array of numbers representing the frequency of occurrence of each pitch class in the generated stimuli representing the Hypophrygian mode ordered along the chromatic scale, i.e., from C<sub>4</sub> to B<sub>4</sub>, I called the event frequency profile for the Hypophrygian mode. The array of numbers representing the frequency of occurrence of each pitch class in the generated stimuli representing the Lydian mode ordered along the chromatic scale, i.e., from C<sub>4</sub> to B<sub>4</sub>, I called the event frequency profile for the Lydian mode. Just like the probe tone profile, the event frequency profile is an array of 12 numbers.

*Stimuli for two-alternative forced-choice task.*

Sixteen additional stimuli were used in Experiment II b: Eight additional Hypophrygian phrases with the least variability in their probe tone profiles (calculated with the data from Experiment I) were selected from the pool of stimuli used in Experiment I (see Table 2). All eight remaining Lydian phrases that were recorded for Experiment I were used.



*Figure 6.* Frequency distribution of pitch classes for the Hypophrygian (A) and Lydian mode (B) displayed as probability of occurrence. The distribution described by Huron and Veltman (HV; 2006) is displayed in black bars. The distribution for the Hypophrygian and Lydian mode as used in my experiment to generate stimuli is displayed in grey bars. The pitch class occurring most in the Hypophrygian mode is F, closely followed by G. The pitch class occurring most in the Lydian mode is C, followed by A. The pitch classes C #, D #, F #, and G # were omitted from the graphs, as they are not used in the Hypophrygian and Lydian mode.

***Procedure.***

Participants gave probe tone ratings to the two chosen probe tone contexts (see Stimuli for probe tone ratings). Afterwards, participants were exposed to 20 minutes of generated stimuli (see Stimuli for exposure phase). The generated stimuli either represented the Hypophrygian or the Lydian mode. Half of the musicians and half of the nonmusicians were exposed to generated stimuli representing the Hypophrygian mode. The other half of the participants was exposed to generated stimuli representing the Lydian mode. Participants had to initiate the start of a new melody by pressing a key. Furthermore, they were asked after each tenth stimulus to indicate how much they liked the past ten stimuli. These measures were introduced to keep participants engaged.

Following exposure, participants gave probe tone ratings to the same two chosen probe tone contexts as prior to exposure (see Stimuli for probe tone ratings). Thus, a total of four probe tone profiles were collected, two prior to exposure (pre-exposure), and two after exposure (post-exposure).

In Experiment II b, after collection of the two post-exposure probe tone profiles, participants completed eight trials of a two-alternative forced-choice task. In this task, a Hypophrygian and a Lydian pre-composed stimulus were paired (see Stimuli for two-alternative forced-choice task). The stimuli paired in a forced-choice trial were selected randomly from the additional pre-composed stimuli included in this experiment. The stimuli were separated by a pause of 1.5 s. After another pause of 1.5 s, the participant was asked to choose the excerpt that seemed more familiar (see Figure 7 for an illustration).

Both experiments took place in a sound attenuated chamber. The experimental programs were written using MATLAB and presented using a Dell Precision T1500 PC. Participants were instructed to adjust the volume to a comfortable level.

pre-composed stimulus, mode 1	1.5 s pause	pre-composed stimulus, mode 2	1.5 s pause	“Which of the two melodies sounded more familiar?”
-------------------------------	-------------	-------------------------------	-------------	--

*Figure 7.* Classification task. Participants are asked which of the two melodies that are played interspersed by pauses of 1.5 s sounded more familiar.

***Data analysis.***

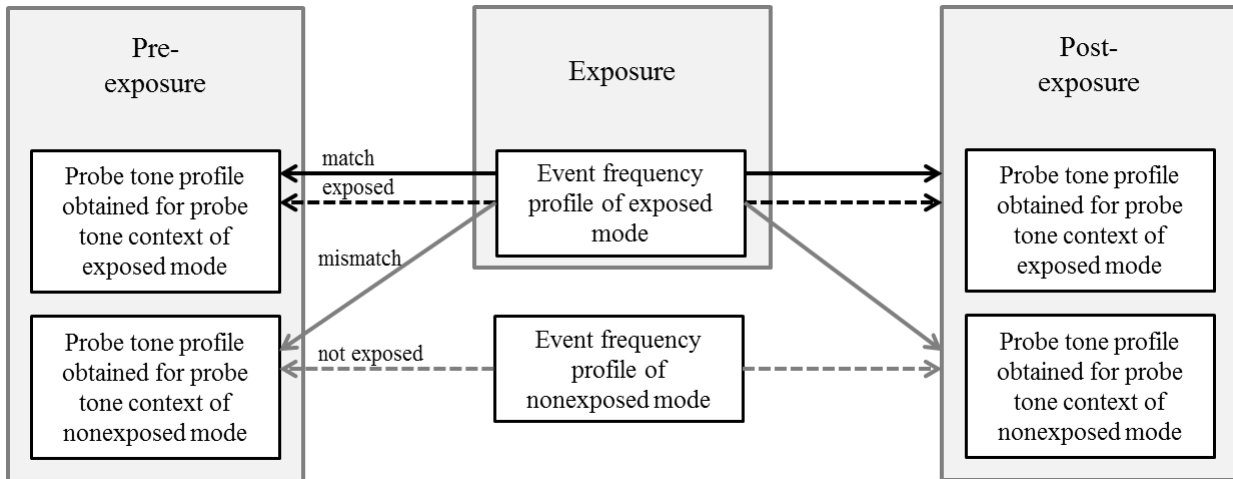
The data from Experiment II a and Experiment II b collected using the same procedure (up until the two-alternative forced-choice task) were analyzed together to increase the power of the analysis. The pre- and post-exposure probe tone profiles of each participant were regressed on the event frequency profiles for the Hypophrygian or Lydian mode (see Stimuli for exposure phase). The beta weights from these regressions were used as dependent variables in two mixed-model ANOVA. The dependent variables can be understood as measures of assimilation, such that higher beta weights indicate more assimilation to the predictor (the event frequency profile of the predicting mode), and lower beta weights indicate less assimilation. See the Appendix for a table summary of how I calculated of my dependent variables, and a visualization of average probe tone profiles. Figure 8 visualizes which variables were entered in the regressions to obtain the dependent variables.

For one mixed-model ANOVA, the event frequency profile of the mode matching the mode of the probe tone context was used as a predictor for the respective probe tone profile to obtain the dependent variables. Thus, the event frequency profile of the Hypophrygian mode was used as a predictor for probe tone profiles collected using the Hypophrygian probe tone context,

and the event frequency profile of the Lydian mode was used as a predictor for probe tone profiles collected using the Lydian probe tone context. This meant that participants had either been exposed to the predicting mode (dashed black arrow in Figure 8), or they had not been exposed to the predicting mode (dashed grey arrow in Figure 8). For this ANOVA, there was therefore a within subjects variable of exposure. The dependent variables in this ANOVA express assimilation to the music system that the context is based on.

For the other mixed-model ANOVA, the event frequency profile of the mode to which the participant was exposed was used as the predictor for all probe tone profiles to obtain the dependent variables. Thus, either the mode of the predictor matched (solid black arrow in Figure 8) or mismatched (solid grey arrow in Figure 8) the mode of the probe tone context for which the probe tone profile was collected. If the participant was exposed to tone sequences generated from the Hypophrygian probability profile, then the event frequency profile of the Hypophrygian mode was used as a predictor for all probe tone profiles. If the participant was exposed to tone sequences generated from the Lydian probability profile, then the event frequency profile of the Lydian mode was used as a predictor for all probe tone profiles. For this ANOVA, there was therefore a within subjects variable of match/mismatch. The dependent variables in this ANOVA express assimilation to the music system that the participant heard during exposure.

For both ANOVA there were four factors. One within subjects variable was either exposure (two levels: exposed or not exposed) or match/mismatch (two levels: match or mismatch), as explained above. The remaining three factors were the same in both ANOVA: A within subjects variable of time (two levels: prior or after exposure), a between subjects variable of group (two levels: musician or nonmusician), and a between subjects variable of experiment (two levels: Experiment II a or Experiment II b).



*Figure 8.* Analysis of Experiment II a and Experiment II b. The boxes with a grey frame illustrate the procedure of the overlapping part of Experiment II a and Experiment II b: Before and after exposure to the stimuli that form the event frequency profile of the exposed mode, probe tone profiles were obtained; one for a probe tone context of the exposed mode, one for a probe tone context of the nonexposed mode. The boxes with a black frame show the variables entered in regressions to obtain the beta weights that were used as dependent variables in two mixed-model ANOVA. The arrows point from the predicting variable to the variable that was regressed. To obtain the dependent variables for the exposed/not exposed analysis, the probe tone profiles were regressed on the event frequency profile of the mode of the probe tone context (indicated by dashed arrows, black for exposed, grey for nonexposed). To obtain the dependent variables for the match/mismatch analysis, the probe tone profiles were regressed on the event frequency profile of the exposed mode (indicated by solid arrows, black for match, grey for mismatch).

## Results.

### *Mixed-model ANOVA.*

For both mixed-model ANOVA, there was no significant main effect of experiment (Experiment II a or Experiment II b), and no significant interaction effect involving experiment ( $ps > .05$ ).

### *Exposed/not exposed.*

The dependent variables in this mixed-model ANOVA were the beta weights from regressions for each participant using the event frequency profile of the mode of the probe tone



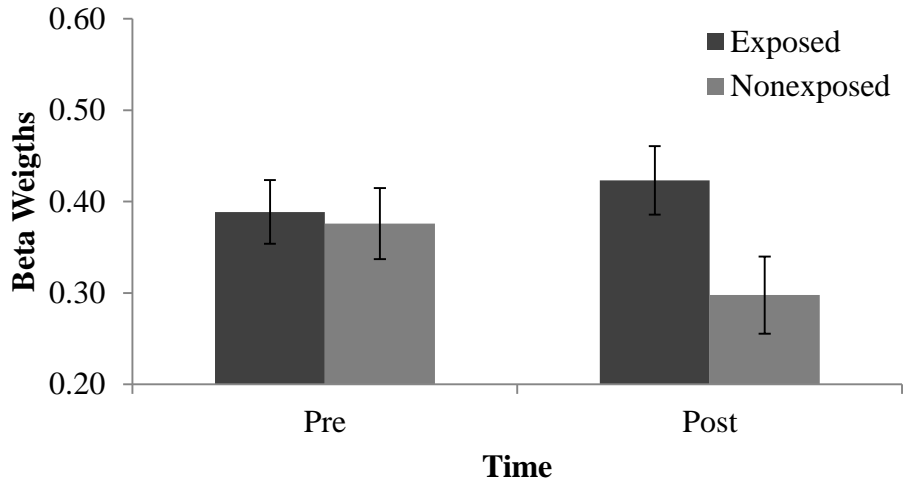
context as a predictor for the probe tone profiles. Thus the participant had either been exposed to the predictor mode or the participants had not been exposed to the predictor mode. The factors of time (within subject), exposure (within subject), and group (between subject) were entered in the analysis. The main effects of time and exposure were not significant (time:  $F(1,39) = 0.70$ ,  $p = .408$ ,  $\eta_p^2 = .02$ ; exposure:  $F(1,39) = 1.27$ ,  $p = .267$ ,  $\eta_p^2 = .03$ ). The main effect of group was significant,  $F(1,39) = 4.76$ ,  $p = .035$ ,  $\eta_p^2 = .11$ . The interaction effect of time and group was marginally significant,  $F(1,39) = 3.93$ ,  $p = .054$ ,  $\eta_p^2 = .09$ . There was a significant interaction of time and exposure,  $F(1,39) = 6.27$ ,  $p = .017$ ,  $\eta_p^2 = .14$ . The interaction is visualized in Figure 9.

Although the three-way interaction between time, group and exposure was nonsignificant,  $F(1,39) = 0.65$ ,  $p > .05$ , paired  $t$ -tests were conducted to investigate which comparisons drove the interaction of time and exposure, and the interaction of time and group. Paired  $t$ -tests on the beta weights of the nonexposed mode for musicians and nonmusicians separately showed no significant change in beta weights for nonmusicians,  $t(19) = 1.01$ ,  $p = .324$ , and a significant decrease in beta weights for musicians,  $t(21) = 2.14$ ,  $p = .044$ . The paired  $t$ -tests on the beta weights of the exposed mode are the same as described later on the beta weights for matching probe tone contexts.

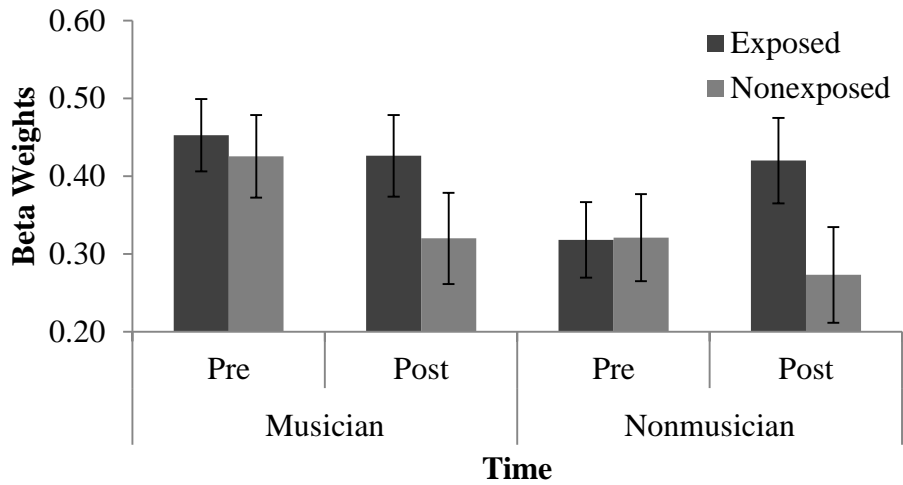
Independent samples  $t$ -tests comparing musicians and nonmusicians on the beta weights of the nonexposed mode revealed no significant difference (pre-exposure:  $t(40) = 1.35$ ,  $p = .183$ ; post-exposure:  $t(40) = 0.55$ ,  $p = .584$ ). The independent samples  $t$ -tests on the beta weights of the exposed mode are the same as described later on the beta weights for matching probe tone contexts.

The BMRQ scores did not correlate with any of the dependent variables in the exposed/not exposed analysis ( $ps > .05$ ).

**A**



**B**



*Figure 9.* Results from exposed/not exposed analysis. The results are shown collapsed across all participants (A) and for musicians and nonmusicians separately (B). Error bars represent standard error of the mean. As can be seen the significant interaction of time and exposure is driven by a decrease for nonexposed beta weights in musicians and an increase for exposed beta weights in nonmusicians.

### *Match/mismatch.*

The dependent variables in this mixed-model ANOVA were the beta weights from regressions for each participant using the event frequency profile of the exposure mode as a predictor for the probe tone profiles. Thus the mode of the probe tone context for which a probe tone profile was collected either matched or mismatched the exposure mode. The factors of time (within subject), match/mismatch (within subject) and group (between subject) were entered in the analysis. There was no significant effect of match/mismatch,  $F(1,39) = 0.00$ ,  $p = .979$ ,  $\eta_p^2 = .00$ . The main effect of time was marginally significant,  $F(1,39) = 3.03$ ,  $p = .090$ ,  $\eta_p^2 = .07$ . The main effect of group was significant,  $F(1,39) = 4.86$ ,  $p = .034$ ,  $\eta_p^2 = .11$ . As can be seen in Figure 10, the beta weights from the analyses using musicians' probe tone profiles were higher than the beta weights from the analyses using nonmusicians' probe tone profiles. The interaction of time and group was not significant at  $F(1,39) = 2.56$ ,  $p = .118$ ,  $\eta_p^2 = .06$ .

Again, there was no significant three-way interaction between time, group and match/mismatch,  $F(1,39) = 0.41$ ,  $p > .05$ . However, to determine the effects that drove the significant effect of group and the marginally significant effect of time paired *t*-tests were conducted. Paired *t*-tests for musicians and nonmusicians separately on the beta weights for matching probe tone contexts showed no significant change in beta weights for musicians,  $t(21) = 0.72$ ,  $p = .482$ , and a significant increase in beta weights for nonmusicians,  $t(19) = 2.11$ ,  $p = .048$ . Paired *t*-tests for musicians and nonmusicians separately on the beta weights for mismatching probe tone contexts showed no significant change in beta weights for musicians,  $t(21) = 0.15$ ,  $p = .879$ , but a trend in beta weights for nonmusicians,  $t(19) = 1.72$ ,  $p = .102$ .

Independent samples *t*-tests comparing musicians and nonmusicians revealed a marginally significant effect of higher beta weights for matching probe tone contexts in

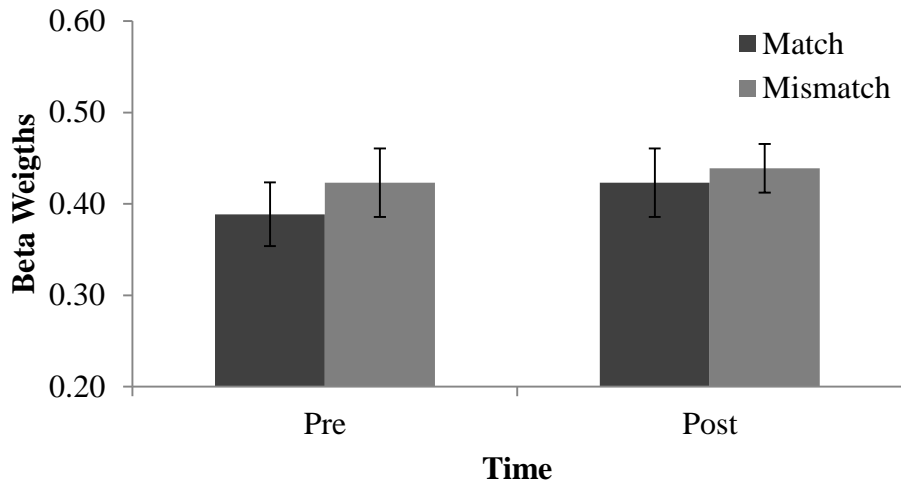
musicians pre-exposure,  $t(40) = 2.00, p = .052$ . Pre-exposure beta weights for mismatching probe tone contexts were not significantly higher in musicians,  $t(40) = 1.35, p = .183$ . The post-exposure beta weights were similar for musicians and nonmusicians (match:  $t(40) = 0.08, p = .936$ ; mismatch:  $t(40) = 0.86, p = .396$ ).

The BMRQ scores did not correlate with any of the dependent variables in the match/mismatch analysis ( $ps > .05$ ).

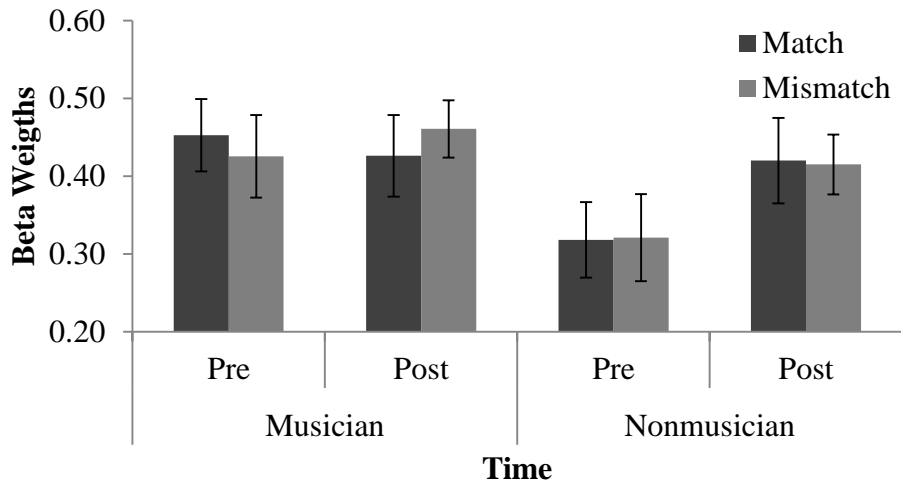
### ***Classification.***

The two-alternative forced-choice trials served as a classification task. Participants' indication of the more familiar excerpt was regarded as a correct classification if participants chose the excerpt representing the mode they were exposed to. A percent correct score was calculated by dividing the number of correct classifications by the number of trials. A one-sample  $t$ -test against .50 (which would indicate chance performance) revealed no significant effect,  $t(21) = 0.34, p = .734$ . For musicians the average percent correct was  $M = .53$  with a standard deviation of  $SD = .23$ . The average percent correct for nonmusicians was  $M = .43$  with a standard deviation of  $SD = .23$ . The BMRQ scores did not correlate with the percent correct score ( $ps > .05$ ).

**A**



**B**



*Figure 10.* Results from match/mismatch analysis. The results are shown collapsed across all participants (A) and for musicians and nonmusicians separately (B). Error bars represent standard error of the mean. As can be seen the marginally significant effect of time was driven by different beta weights pre- and post-exposure in nonmusicians. There is no effect of match/mismatch.

## **Discussion.**

The analyses showed that musicians and nonmusicians behave differently: The beta weights were generally higher for musicians than nonmusicians, as evidenced by the significant main effects of group in both the match/mismatch and the exposed/not exposed analysis.

The exposed/not exposed analysis revealed an interaction of time and exposure. The beta weights decreased for the nonexposed mode (especially for musicians), and increased for the exposed mode (especially for nonmusicians). Thus, the event frequency profile of the exposed mode gained predictive power, whereas the event frequency profile of the nonexposed mode lost predictive power. This suggests that participants were able to abstract the statistical regularities in the generated tone sequences that were presented to them during exposure.

The match/mismatch analysis investigates the data from a more participant driven view (as opposed to the more predictor driven exposed/not exposed analysis): The marginally significant main effect of time in the match/mismatch analysis also suggests that musicians and nonmusicians abstracted the event frequency profile put forth during exposure. However, looking at musicians and nonmusicians separately, it appears that musicians' mental hierarchy of the novel musical system was resistant to influence of exposure. Nonmusicians' mental hierarchy of the novel musical system on the other hand displayed influence of exposure, such that the prediction of the probe tone profiles using the event frequency profile of the exposure mode improved after exposure. Therefore, the marginally significant main effect of time in the match/mismatch analysis was driven by the significant difference between beta weights prior to and after exposure in nonmusicians.

The differential change is especially visible in Figure 10: While nonmusicians' beta weights increase, the musicians' beta weights post-exposure remain similar to the beta weights

pre-exposure. Paired *t*-tests confirm that the beta weights for musicians remain similar, and that the beta weights of matching probe tone contexts post-exposure for nonmusicians are higher than prior to exposure. Moreover, after exposure, nonmusicians' beta weights were at a similar level as musicians' beta weights, whereas prior to exposure musicians' beta weights were higher for matching probe tone contexts.

Thus, the significant effect of group in the match/mismatch analysis was driven by a significant difference in beta weights between musicians and nonmusicians prior to exposure. It could be argued then, that nonmusicians, i.e., participants with little music training, are the better sample to investigate unintentional learning of pitch hierarchy, because the mental hierarchy of pitch in musicians appears to be resistant to change.

Results from the match/mismatch analysis suggest that participants treated probe tone contexts of different modes in the same way, as there was no significant effect involving the factor match/mismatch. Although it appears that at least nonmusicians learned during exposure, they did not treat the probe tone context that was not of the same mode as the tone sequences heard during exposure any differently than the probe tone context that was of the same mode as the tone sequences heard during exposure. This is also evidenced by the results from the classification task: Participants (musicians and nonmusicians) did not behave different from chance when indicating which stimulus they found more familiar.

Overall, unintentional learning of novel pitch hierarchies appears to be possible after 20 min of exposure. Nonmusicians especially may be a well suited sample to investigate unintentional learning of novel pitch hierarchy. However, participants treated probe tone contexts representing different pitch hierarchies in the same way.

Why were participants unable to distinguish between modes? Perhaps the exposure time was not sufficient, so that participants were able to abstract the statistical regularities to a certain degree, but not to such an extent that participants were able to distinguish it from another probability profile.

To keep the experiment short, participants in my experiments were listening to melodies for 20 min, whereas, for instance, the exposure in a study by Loui et al. (2010) that investigated passive statistical learning of a novel music system based on second-order probabilities was 30 minutes long. Loui et al. (2010) reported successful generalization of the probabilities after 30 minutes in a classification task similar to mine, and no differences between people with more music training and people without any music training. The group with more music training in the study by Loui et al. (2010) however, was only required to have more than five years of music training, which might have resulted in a less trained group of participants than the group of people who were classified as musicians in my experiment.

However, it should be noted, that Loui et al. (2010) investigated learning of transitional rules, i.e. a set of second-order probabilities. It could be argued that learning of a second-order probability system is easier than learning of a first-order probability system, or at least easier to show. As soon as an illegal transition from one pitch to another is made, i.e., a transition that is not possible given the second-order probability rules), the participant can indicate that the tone sequence is unfamiliar. However, in first-order probability systems, like the pitch hierarchy of the Hypophrygian or Lydian mode as used in my experiments, there are few pitches that allow the same kind of definitive classification. If learning of a first-order probability system is harder, perhaps the exposure time needs to be extended to longer than 30 minutes.



### Chapter 3: General Discussion

The sensitivity to statistical information in tone sequences of a novel musical system as quantified by the probe tone method seems to differ between musicians and nonmusicians. A significant correlation between years of training and my measure of general sensitivity in Experiment I suggests that sensitivity is heightened with more music training. More experienced musicians might be more careful listeners if presented with musical stimuli, or they may be more adept at picking up statistical information quickly. Kraus and Chandrasekaran (2010) argued that musicians' more adaptive auditory systems are trained to pick out relevant sounds from soundscapes, and gives rise to improved ability to track regularities. In this experiment, those could be recurring, i.e., salient pitch classes. Musicians could also be more aware of the importance of salient pitch classes due to their extensive training.

Another factor that could have aided musicians in this task is better working memory (Pallesen et al., 2010). Better working memory would mean that musicians are potentially able to consider more of the probe tone context when making goodness-of-fit judgments. It should be noted that despite having received little to no music training, nonmusicians exhibited a marginally significant correlation between their probe tone profiles and the event frequency distribution of the probe tone contexts. This implies that tallying up the pitch class occurrences is a strategy employed by nonmusicians and musicians alike.

In light of these results, a difference between musicians and nonmusicians in Experiments II a and Experiment II b was expected. The results conformed to this expectation: Musicians seemed resistant to effects of exposure, whereas nonmusicians abstracted the frequency information contained in the exposure melodies.

The post-exposure probe tone profiles compared to the pre-exposure probe tone profiles of nonmusicians showed that they gained musical knowledge over the course of exposure. They generalized the statistical information contained in the exposure melodies to two probe tone contexts, though one of the contexts did not match the exposure mode. It seems as if nonmusicians learned the pitch hierarchy, but did not distinguish it from the pitch hierarchy of another mode. This is echoed in the results from the classification task in Experiment II b. The percent correct of nonmusicians' was not different from chance.

Based on the percent correct score from the classification task in Experiment II b, it appears that musicians did not distinguish between modes either. Furthermore, it appears that musicians were also not susceptible to effects of exposure, as the measures of assimilation to the pitch hierarchy (the beta weights) remained unchanged after exposure. This difference suggests that training leads to an internal representation of musical structure that is less likely to be influenced by short exposure to a novel musical system.

This gives rise to the question of whether exposure time in my experiments should be extended. As mentioned previously, the exposure time in Experiments II a and Experiment II b was less than the exposure time in studies conducted in other labs that concerned learning of a novel music system (Loui & Wessel, 2008; Loui et al., 2010). These studies involved participants being exposed to 30 minutes of musical stimuli that were based on a second-order probability system.

In the study by Loui et al. (2010), contrary to my experiments, there were no differences between participants with more training and participants with little or no training. However, the participants in the study by Loui et al. (2010) with more music training might be less trained than the participants who were classified as musicians in my study. Thus, it could be argued, that

nonmusicians, i.e., participants with little or no music training, are the better suited sample to investigate learning of statistical regularities in music.

Bigand and Poulin-Charronnat (2006) have proposed that many differences found between musicians and nonmusicians can be attributed to methodology that favors “correct” responses by musicians, i.e., if technical musical terms are included in experimental instructions. However, I believe that my experimental instructions were easy enough to understand for nonmusicians by using technical terms that are used in everyday language (“tone” and “melody” in “How well does the single tone fit in the melody played before?”, “melody” in “Which of the two melodies sounded more familiar?”). Thus, the higher sensitivity to frequency information by musicians compared to nonmusicians in Experiment I and, prior to exposure, in Experiment II a and Experiment II b, as well as the correlation between years of training and my measure of general sensitivity in Experiment I, suggest differences are due to formal music training. As I mentioned previously, music training might have led to better working memory and a more adaptive auditory system which could have helped musicians display higher sensitivity (Kraus & Chandrasekaran, 2010; Pallesen et al., 2010; Parbery-Clark, Skoe, Lam & Kraus, 2009).

In a study by Bigand, D’Adamo, and Poulin-Charronnat (cited in Bigand & Poulin-Charronnat, 2006), musically trained and untrained listeners were exposed to stimuli generated from the same serial order of pitches. In an ensuing classification task a tone sequence generated from the same serial order of pitches, and one generated from another serial order of pitches, were paired. Participants had to indicate which of the two stimuli was composed in the same way as those heard during the exposure phase. Bigand et al. (cited in Bigand & Poulin-Charronnat, 2006) found no differences between musically trained and untrained listeners. Both groups of listeners performed above chance. However, the classification of musically trained versus

untrained listeners was not specified. As with the sample used in the study by Loui et al. (2010), the musically trained listeners could have had less extensive training than the participants who I classified as musicians in my study. It should also be noted that the serial order of pitches is a set of transitional rules.

In conclusion, it can be said that participants made use of the statistical information that was contained in the stimuli. In Experiment I, participants had no other information about the style-appropriate tonal hierarchy of the probe tone contexts other than the event frequency distribution of the probe tone context. The latter correlated with the goodness-of-fit ratings. Assuming the probe tone profile represents the mental pitch hierarchy, this finding indicates that participants learned the pitch hierarchy put forth in the musical stimulus.

In Experiments II a and II b, participants had no additional information about the style-appropriate pitch hierarchy of the mode of the probe tone contexts prior to exposure. The event frequency profile of the exposure mode correlated with goodness-of-fit ratings. In order for the event frequency profile of the exposure mode to influence the goodness-of-fit ratings, participants had to abstract it. This also indicates that participants learned the pitch hierarchy put forth in the musical stimuli. In contrast to Experiment I, this learning is also evidenced in a change in assimilation for nonmusicians. Thus, the pitch hierarchy was updated during exposure to incorporate the statistical information contained in the melodies during exposure.

Altogether, these findings provide evidence for the third and last proposition of the Theory of Tonal Hierarchies in Music (Krumhansl & Cuddy, 2010) which states that “listeners rapidly adapt to style-appropriate tonal hierarchies even if the style is unfamiliar” (p. 80).

My experiments also add to the research of unintentional learning by extending research to passive statistical learning of first-order probabilities. This extension seemed overdue as the

original research, which led to psychologists suggesting that listeners unintentionally learn about music through exposure, was on the mental representation of a first-order probability system (the tonal hierarchy, see Castellano, Bharucha, & Krumhansl, 1984, and responses by Bharucha, 1984, and Deutsch, 1984).

There are two directions in which I want to extend my research:

First, by extending or shortening the exposure time, I want to investigate the effect of exposure in more detail. Would shorter exposure time lead to even less change in beta weights and longer exposure time to more change? This line of investigation also opens up possibilities to introduce other dimensions to this research: For instance, by assessing participants' appraisal of the novel musical system after exposure I could investigate the mere exposure effect. The mere exposure effect predicts "emerging preference" (page 224) to a stimulus through exposure to it (Zajonc, 2001).

Loui et al. (2010) dispute that there is a relationship between knowing music and liking it (preference). On the other hand, a recent study by Szpunar et al. (2004) described the relationship between knowing and liking music as curvilinear, following an inverted U-shape. However, the relationship is linear (as predicted by the mere exposure effect) if participants are listening to music incidentally (Schellenberg et al., 2008). Thus, these authors posit that exposure and preference are linked. It should be noted, however, that both Szpunar et al. (2004) and Schellenberg et al. (2008) reached their conclusions using excerpts from classical music pieces as stimuli. Using the same paradigm as in my experiments would provide: a) A more controlled design, as the musical stimuli are novel and thus participants are highly unlikely to have encountered them before; and b) a more detailed assessment of participants' knowledge of music (the pitch hierarchy) than a pure recognition task as employed by the two referenced studies.

Second, by including neuropsychological measures, I could further explore the differences between musicians and nonmusicians. As music training increases MMN amplitudes in children already (Putkinen et al., 2014), it could be assumed that this pattern might persist later in development.

The MMN especially could be an interesting negativity to study with the paradigm used in my experiments. The MMN is thought to reflect departures from regularities established internally in short-term format, based on the “information buffered in the auditory sensory memory” (Koelsch, 2013, page 55). Probe tones whose pitch classes occurred in the probe tone context could be interpreted as standard stimuli, whereas probe tones whose pitch classes did not occur in the probe tone context could be used as oddball stimuli.

If participants use the event frequency count of the probe tone context to establish musical regularities, then the oddball stimuli should elicit a negativity that could be characterized as MMN (Koelsch, Schröger, & Tervaniemi, 1999; Näätänen & Picton, 1987; Rammsayer & Altenmüller, 2006; Saarinen, Paavilainen, Schröger, Tervaniemi, & Näätänen, 1992; Sams, Paavilainen, Alho, & Näätänen, 1985; Tervaniemi, Ilvonen, Karma, Alho, & Näätänen, 1997).

If, however, they establish those regularities in long-term format, the negativity might be characterized better as an ERAN. Oddball stimuli are acoustically uncommon events if regularities are not established in long-term format. Oddball stimuli become structurally uncommon events after regularities have been established in long-term format (Koelsch et al., 2007).

As Salimpoor et al. (2013) pointed out, ERP measures might be more sensitive than behavioral paradigms in reflecting passive learning. Furthermore, by including neuropsychological measures in the paradigm that was used I could pinpoint what exactly was

learned: If MMN amplitudes become more pronounced after exposure, it could indicate that participants paid more attention and thus exhibited enhanced auditory discrimination, whereas if an ERAN develops after exposure, participants probably established regularities in long-term format, i.e., they truly learned.

The possibility that increased beta weights did not reflect abstraction of statistical regularities but rather enhanced auditory discrimination presents a limitation to the interpretation of my results. If the MMN amplitudes become more pronounced after exposure, but no ERAN develops, the interpretation of previous research on tonal hierarchies that attributed the more pronounced probe tone profiles by participants with more music training to the experience of those participants would also have to be re-evaluated. Including neuropsychological measures would help to shed light on the exact nature of improvement in beta weights. However, adding neuropsychological measures would come with its own set of unique methodological problems. As ERP measures are sensitive to movements by the participant for instance, ERPs would have to be collected for a number of trials. As a probe tone profile consists of 12 probe tone ratings, which would have to be collected repeatedly, the length of the experiment would increase substantially.

In future studies I would also like to rethink the stimuli. Using Hypophrygian and Lydian chants guarantees a certain degree of musicality (as they were once everyday music). However, creating a new, unique probability profile would facilitate control over the frequency count of pitch classes in my stimuli. By creating my own probability profiles I could therefore constrict the number of pitch classes that occur, which would make the definition of potential oddball stimuli easier.

Overall, my experiments constituted the starting point for research of abstraction of first-order probabilities, from which I can venture further. My experiments broaden our knowledge of statistical learning by expanding existing research to include abstractions of first-order probabilities. While learning of frequencies of occurrence in music has been so far unexplored (Rohrmeier & Rebuschat, 2012), recently, the learning of frequencies of occurrence in language has been discussed (Werker, Yeung, & Yoshida, 2012). This indicates that the abstraction of first-order probabilities, not unlike the abstraction of second-order probabilities, may be a domain general cognitive process. There is much left to be explored.



## References

- The Benedictines of Solesmes (Eds.) (1961). *The Liber Usualis with Introduction and Rubrics in English*. Tournai: Desclée Company.
- Bharucha, J. J. (1984). Event hierarchies, tonal hierarchies, and assimilation: A reply to Deutsch and Dowling. *Journal of Experimental Psychology: General*, *113*(3), 421–425.
- Bigand, E., D’Adamo, D. A., & Poulin-Charronnat, B. The implicit learning of twelve-tone music. Cited in 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*(1), 115–117.
- 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*(1), 100–130.
- Burgoyne, J. A., Wild, J., & Fujinaga, I. (2013). Compositional data analysis of harmonic structures in popular music. In *Mathematics and Computation in Music* (pp. 52–63). Berlin: Springer.
- Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, *113*(3), 394–412.
- Collett, M. (2013). *Acquisition and generalization of pitch probability profiles* (Unpublished master’s thesis). Queen’s University, Kingston.
- Common Scale Types (n.d.). In *Encyclopædia Britannica*. Retrieved from <http://www.britannica.com/EBchecked/topic/526314/scale/64512/Common-scale-types>
- Corrigall, K. A., & Trainor, L. J. (2010). Musical enculturation in preschool children: Acquisition of key and harmonic knowledge. *Music Perception*, *28*(2), 195–200.
- Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception & Psychophysics*, *41*(6), 609–620.
- Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharucha, and Krumhansl. *Journal of Experimental Psychology: General*, *113*(3), 413–416.
- Diatonic (n.d.). In *Encyclopædia Britannica*. Retrieved from <http://www.britannica.com/EBchecked/topic/161859/diatonic>.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458.

- Fritts, L. (2013). *University of Iowa Electronic Music Studios Musical Instrument Samples*. Retrieved from: <http://theremin.music.uiowa.edu/MISpiano.html>.
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, 78(3), B53–B64.
- Huron, D., & Veltman, J. (2006). A cognitive approach to medieval mode: Evidence for an historical antecedent to the major/minor system. *Empirical Musicology Review*, 1, 33–55.
- Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *Neuroimage*, 47(2), 735–744.
- Kessler, E. J., Hansen, C. & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception*, 2, 131–165.
- Knopoff, L., & Hutchinson, W. (1983). Entropy as a measure of style: the influence of sample length. *Journal of Music Theory*, 27, 75–97.
- Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, 46(1), 179–190.
- Koelsch, S. (2013). *Brain and Music*. Hoboken, NJ: John Wiley & Sons, Ltd.
- Koelsch, S., Jentschke, S., Sammler, D., & Mietschen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, 44(3), 476–490.
- Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, 39(5), 657–663.
- Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, 10(6), 1309–1313.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), 599–605.
- Krumhansl, C. L. (1985). Perceiving tonal structure in music. *American Scientist*, 73, 371–378.
- Krumhansl, C. L. (1987). Tonal and harmonic hierarchies. In J. Sundberg (Ed.), *Harmony and tonality* (pp. 13–32). Stockholm, Sweden: Royal Swedish Academy.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York, NY: Oxford University Press.
- Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (pp. 51–87). New York: Springer.

- 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, 334–368.
- Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a tonal context. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 579–594.
- Lantz, M. E., Kim, J.-K. & Cuddy, L. L.. Perception of a tonal hierarchy derived from Korean music. *Psychology of Music*, 0(0), 1–19.
- Loui, P., & Wessel, D. (2008). Learning and liking an artificial musical system: Effects of set size and repeated exposure. *Musicae Scientiae*, 12(2), 207–230.
- 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.
- Loui, P. & Schlaug, G. (2012). Impaired learning of event frequencies in tone deafness. *Annals of the New York Academy of Sciences*, 1252(1), 354–360.
- Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J., & Rodriguez-Fornells, A. (2013). Individual differences in music reward experiences. *Music Perception*, 31(2), 118–138.
- Meyer, L.B. (1957). Meaning in music and information theory. *The Journal of Aesthetics and Art Criticism*, 15(4), 412–424.
- Miller, G. A., & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *The American journal of psychology*, 176–185.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4), 375–425.
- Oram, N., & Cuddy, L. L. (1995). Responsiveness of western adults to pitch distributional information in melodic sequences. *Psychological Research*, 57, 103–118.
- Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PloS one*, 5(6), e11120.
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and hearing*, 30(6), 653–661.
- Peretz, I., Gaudreau, D. & Bonnel, A. (1998). Exposure effects on music preference and recognition. *Memory & Cognition*, 26(5), 884–902.
- Putkinen, V., Tervaniemi, M., Saarikivi, K., Vent, N. D., & Huotilainen, M. (2014). Investigating the Effects of Musical Training on Functional Brain Development with a Novel Melodic MMN Paradigm. *Neurobiology of learning and memory*, 110, 8–15.

- Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception, 24*(1), 37–48.
- Rohrmeier, M. & Rebuschat, P. (2012). Implicit learning and acquisition of music. *Topics in Cognitive Science, 4*, 525–553.
- Saarinen, J., Paavilainen, P., Schöger, E., Tervaniemi, M., & Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport, 3*(12), 1149–1151.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*(2), 101–105.
- Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science, 340*(6129), 216–219.
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 62*(6), 437–448.
- Schellenberg, E. G., Peretz, I., & Vieillard, S. (2008). Liking for happy-and sad-sounding music: Effects of exposure. *Cognition & Emotion, 22*(2), 218–237.
- Smith, N. A., & Schmuckler, M. A. (2004). The perception of tonal structure through the differentiation and organization of pitches. *Journal of Experimental Psychology: Human Perception and Performance, 30*(2), 268–286.
- Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 370–381.
- Temperley, D. (2007). *Music and Probability*. Cambridge, Mass: MIT Press.
- Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., & Näätänen, R. (1997). The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience letters, 226*(1), 1–4.
- Tillmann, B., & McAdams, S. (2004). Implicit learning of musical timbre sequences: statistical regularities confronted with acoustical (dis) similarities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(5), 1131–1142.
- Thiessen, E. D., & Erickson, L. C. (2013). Beyond word segmentation: A two-process account of statistical learning. *Current Directions in Psychological Science, 22*(3), 239–243.

- Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception and Psychophysics*, *56*(2), 125–132.
- Trainor, L. J., Marie, C., Gerry, D., Whiskin, E., & Unrau, A. (2012). Becoming musically enculturated: effects of music classes for infants on brain and behavior. *Annals of the New York Academy of Sciences*, *1252*(1), 129–138.
- Werker, J. F., Yeung, H. H., & Yoshida, K. A. (2012). How do infants become experts at native-speech perception?. *Current Directions in Psychological Science*, *21*(4), 221–226.
- Zajonc, R.B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, *10*(6), 224–228.
- Zatorre, R. J., & Salimpoor, V. N. (2013). From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences*, *110* (Supplement 2), 10430–10437.

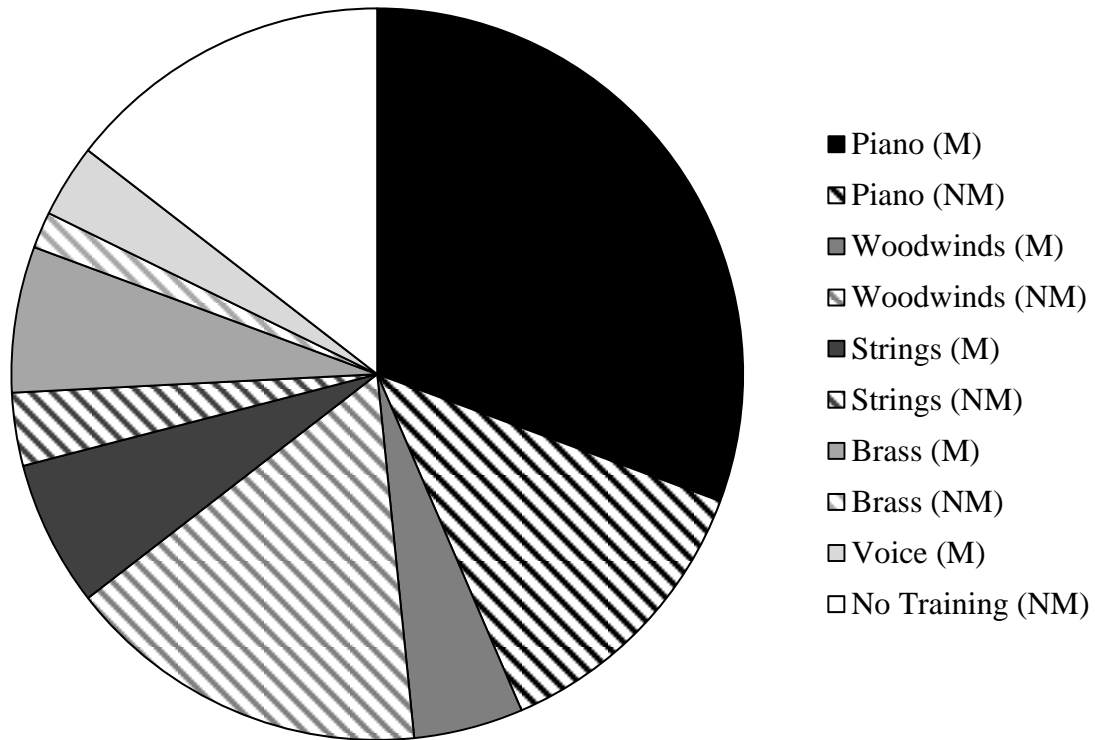
### Appendix I: Music Training Related Descriptors of Sample

*Average age at start of training for musicians (M) and nonmusicians (NM) by experiment. Independent samples t-tests were calculated for each experiment to determine if there was a significant difference between musicians and nonmusicians. Only participants who had had music training were considered for this table.*

<b>Experiment</b>	<b>M (M)</b>	<b>SD (M)</b>	<b>M (NM)</b>	<b>SD (NM)</b>	<b>Independent samples t-test</b>
I	6.90	3.18	11.80	1.48	$t(13) = 2.23, p = .007$
II a	7.45	4.04	10.14	2.85	$t(15) = 1.51, p = .151$
II b	8.58	3.15	11.33	3.43	$t(19) = 1.91, p = .072$
all	7.70	3.42	11.05	2.84	$t(51) = 3.71, p = .001$

*Type of certification for musicians. Participants were classified as musicians, if they held at least Grade X RCM certificates or had taken university level music classes. The Associate of the Royal Conservatory of Music (ARCT) diploma is the highest academic credential awarded by the RCM.*

<b>Type of certification</b>	<b>Number of musicians</b>
Grade X RCM certificate or equivalent	16
ARCT	4
University level music classes	12



Distribution of instrumental and vocal training. Musicians are depicted as solid segments; nonmusicians are depicted as ruled segments. Almost two thirds of the musicians were trained on the piano (19 of 32 trained on piano, 4 each on string and brass instruments, 3 on woodwind and 2 on voice). Nonmusicians were mostly trained on woodwind instruments or piano (10 on woodwind instruments, 8 on piano, 2 on string and 1 on brass instruments, 9 participants without any instrumental or vocal training).

## Appendix II: Descriptors of Sample Not Related to Music Training

Forty participants were female, 22 male.

English was the first language of 46 participants. Five musicians and two nonmusicians reported Chinese as first language; there was one musician each reporting German, Korean and Tagalog as first language, two nonmusicians each reporting French and Spanish as first language and one nonmusician each reporting Korean and Serbo-Croatian as first language. 58 of the participants were right handed; 1 musician and 3 nonmusicians were left handed.

All participants reported normal hearing. 5 musicians and 1 nonmusician reported perfect pitch; 1 musician reported relative pitch.

Participants were 21.54 years old on average ( $SD = 3.09$ ). Musicians were 20.79 years old on average ( $SD = 3.40$ ). Nonmusicians were 22.23 years old on average ( $SD = 2.66$ ).

Participants were also asked to indicate their favorite genre. None of the participants listed medieval or Gregorian chant music as their favorite genre. The following table lists the genres that were listed as favorite genres by group (either musicians or nonmusicians).

*Favorite genre by musicians (M) and nonmusicians (NM). Genres that were listed twice or less were counted under "others", these included country, electronic, and R&B.*

<b>Genre</b>	<b>M</b>	<b>NM</b>
Acoustic	2	1
Classical	12	0
Jazz	2	1
Pop	5	6
Rap	1	2
Rock	9	14
Others	1	3



The BMRQ was administered to participants in Experiment II a and Experiment II b. The BMRQ assesses participants on their use of and relationship with music. Five factors (Music Seeking, Emotion Evocation, Mood Regulation, Social Reward and Sensory-Motor) are identified. Their average score and standard deviation next to the average score and standard deviation of the overall questionnaire (Music Reward) are listed in Table 6.

The scores were computed using the online calculator provided by the authors (<http://www.brainvitge.org/bmrq.php>). There were 41 completed questionnaires. Unlike the results presented by Mas-Herrero et al. (2013), differences between musicians and nonmusicians were not found on the factors Music Seeking and Emotion Evocation. However a difference was found on the factor Social Reward as reported by Mas-Herrero et al. (2013)

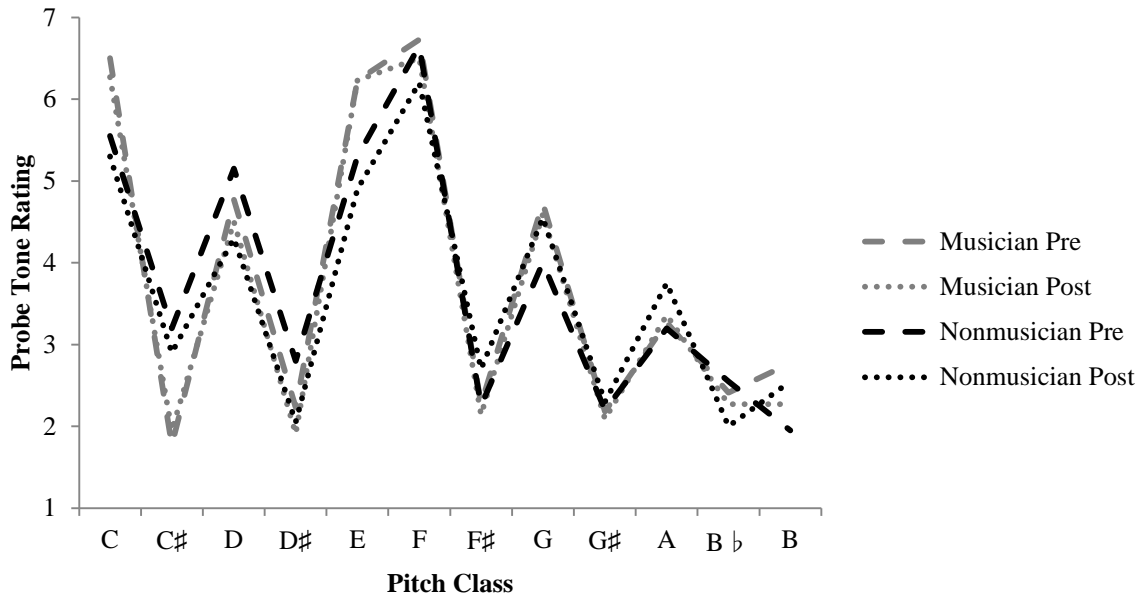
*BMRQ scores for musicians (M) and nonmusicians (NM). Independent samples t-tests were calculated for each experiment to determine if there was a significant difference between musicians and nonmusicians.*

<b>Scale</b>	<b>M (M, SD)</b>	<b>NM (M, SD)</b>	<b>Independent samples t-test</b>
Music Seeking	55.50, 9.19	51.16, 9.25	$t(39) = 1.50, p = .141$
Emotion Evocation	51.77, 9.31	49.94, 9.90	$t(39) = 0.61, p = .547$
Mood Regulation	51.23, 9.06	48.11, 14.64	$t(39) = 0.83, p = .410$
Sensory-Motor	45.82, 11.75	49.16, 10.64	$t(39) = 0.95, p = .349$
Social Reward	57.73, 8.86	49.47, 12.04	$t(39) = 2.52, p = .016$
Music Reward	53.32, 10.17	49.58, 12.26	$t(39) = 1.07, p = .256$

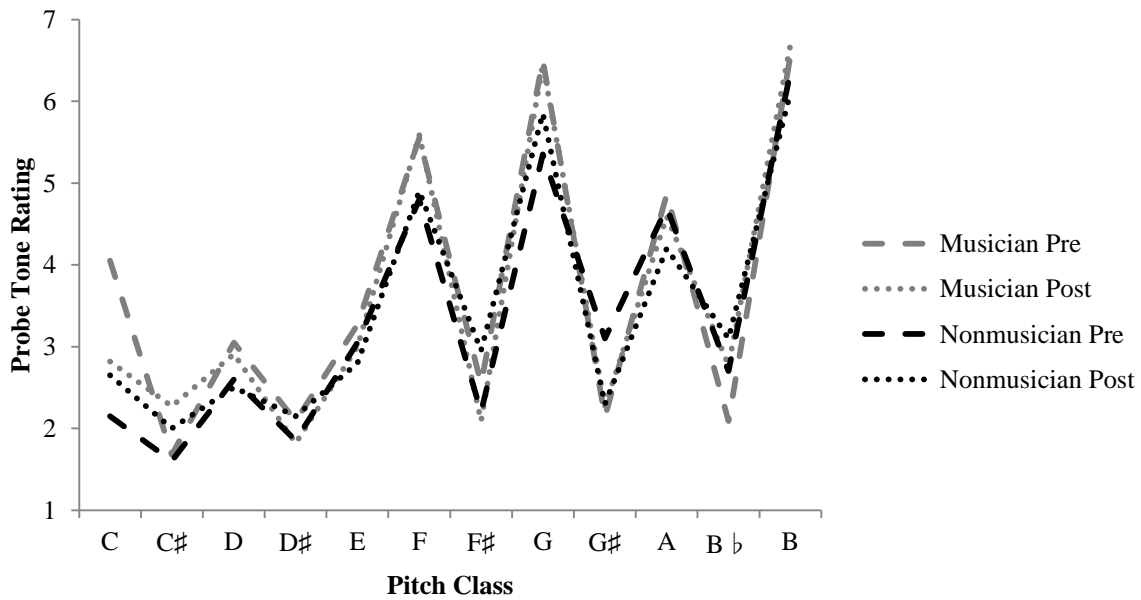
### Appendix III: Calculation of Dependent Variables and Average Probe Tone Profiles

*Variables entered in regressions to obtain dependent variables.*

<b>Mode of exposure tone sequence</b>	<b>Mode of predicting event frequency profile (predictor in regression)</b>	<b>Mode of probe tone profile (dependent variable in regression)</b>	<b>Terminology of dependent variable entered in ANOVA</b>
Hypophrygian	Hypophrygian	Hypophrygian	$\beta$ – matched, exposed
Hypophrygian	Lydian	Hypophrygian	$\beta$ – nonexposed
Hypophrygian	Hypophrygian	Lydian	$\beta$ – mismatched
Lydian	Lydian	Lydian	$\beta$ – matched, exposed
Lydian	Hypophrygian	Lydian	$\beta$ – nonexposed
Lydian	Lydian	Hypophrygian	$\beta$ – mismatched



Average probe tone ratings for Hypophrygian context in Experiment II a and Experiment II b. The mean probe tone rating for each pitch is depicted separately for musicians (grey) and nonmusicians (black) prior (dashed) and after exposure (dotted).



Average probe tone ratings for Lydian context in Experiment II a and Experiment II b. The mean probe tone rating for each pitch is depicted separately for musicians (grey) and nonmusicians (black) prior (dashed) and after exposure (dotted).

## Appendix IV: Research Ethics Approval



October 29, 2013

Ms. Anja-Xiaoxing Cui  
Master's Student  
Department of Psychology  
Queen's University  
Kingston, ON K7L 3N6

**GREB Ref #: GPSYC-636-13; Romeo # 6011156**  
**Title: "GPSYC-636-13 Musical Knowledge"**

Dear Ms. Cui:

The General Research Ethics Board (GREB), by means of a delegated board review, has cleared your proposal entitled "GPSYC-636-13 Musical Knowledge" for ethical compliance with the Tri-Council Guidelines (TCPS) and Queen's ethics policies. In accordance with the Tri-Council Guidelines (article D.1.6) and Senate Terms of Reference (article G), your project has been cleared for one year. At the end of each year, the GREB will ask if your project has been completed and if not, what changes have occurred or will occur in the next year.

You are reminded of your obligation to advise the GREB, with a copy to your unit REB, of any adverse event(s) that occur during this one year period (access this form at [https://eservices.queensu.ca/romeo\\_researcher/](https://eservices.queensu.ca/romeo_researcher/) and click Events - GREB Adverse Event Report). An adverse event includes, but is not limited to, a complaint, a change or unexpected event that alters the level of risk for the researcher or participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GREB within 48 hours.

You are also reminded that all changes that might affect human participants must be cleared by the GREB. For example you must report changes to the level of risk, applicant characteristics, and implementation of new procedures. To make an amendment, access the application at [https://eservices.queensu.ca/romeo\\_researcher/](https://eservices.queensu.ca/romeo_researcher/) and click Events - GREB Amendment to Approved Study Form. These changes will automatically be sent to the Ethics Coordinator, Gail Irving, at the Office of Research Services or [irvingg@queensu.ca](mailto:irvingg@queensu.ca) for further review and clearance by the GREB or GREB Chair.

On behalf of the General Research Ethics Board, I wish you continued success in your research.

Yours sincerely,

A handwritten signature in black ink that reads "Joan Stevenson".

Joan Stevenson, Ph.D.  
Chair  
General Research Ethics Board

c: Dr. Lola Cuddy and Dr. Niko Troje, Faculty Supervisors  
Dr. Stanka Fitneva, Chair, Unit REB  
Ms. Marie Tooley, Dept. Admin.