Acquiring New Musical Grammars: a Statistical Learning Approach

Psyche Loui (psyche@berkeley.edu)
Department of Psychology & Center for New Music and Audio Technologies
University of California at Berkeley, 1750 Arch St, Berkeley, CA, USA 94709

David Wessel (wessel@cnmat.berkeley.edu)
Department of Music & Center for New Music and Audio Technologies
University of California at Berkeley, 1750 Arch St, Berkeley, CA, USA 94709

Carla Hudson Kam (clhudson@berkeley.edu)
Department of Psychology, University of California at Berkeley
3210 Tolman Hall, Berkeley, CA, USA 94709

Abstract

In the present study we examine the ability of humans to acquire knowledge via passive exposure to a new musical system. We designed two new musical grammars based on a non-Western tuning system, and created melodies as legal exemplars of each grammar. In two experiments each participant was exposed to a set of melodies from one grammar. Several tests were conducted to assess learning, including forced-choice recognition and generalization, pre- and post-exposure probe tone ratings, and subjective preference ratings. In Experiment 1, five melodies were presented repeatedly. Participants correctly recognized and preferred melodies they had heard, but failed to generalize their recognition to new exemplars of the same grammar. In Experiment 2, 15 melodies were presented repeatedly. Participants showed some tendency to make generalizations about new melodies in their given grammar, and also showed an increased sensitivity to the statistics of the musical grammar following exposure. Results suggest that larger sets of exemplars promote the extraction of regularities underlying the examples, whereas smaller sets lead to better recognition and are more likely to influence subjective preference.

Introduction

The ability of the human cognitive system to learn the statistical properties of the environment has been posited as a domain-general mechanism subserving human abilities including language (Saffran et al, 1996), motor movements (Hunt & Aslin, 2001), and music (Saffran et al, 1999). In particular, evidence for knowledge in musical harmony has repeatedly been demonstrated in humans through behavioral and physiological methods (Besson & Faïta, 1995; Krumhansl & Kessler, 1983). As many of these findings are robust even in humans without explicit musical training (e.g. Bharucha & Stoeckig, 1986; Koelsch et al, 2000), it would seem that explicit training in music is not required for the acquisition of knowledge in musical harmony, and that this knowledge could be incidentally acquired by humans via passive exposure to music.[1]

A New Paradigm for Studying Musical Acquisition

In the present study we examine this possibility by investigating the learning of artificial musical grammars, with the specific intention of probing for recognition, expectation, and preference formation in the new musical system. We designed two new musical grammars based on a microtonal tuning system, very different from traditional Western music. The musical grammars were constructed such that they obeyed some psychoacoustic and Gestalt principles, but were completely novel to all of our participants. We presented adult participants with melodies generated according to the grammars, and examined their learning of the principles underlying those melodies. To presage the results, we found extensive learning within a relatively short time, supporting the notion that principles of musical harmony can be implicitly acquired by humans via passive exposure.

A New Musical System

Extensive research has been done with common-practice Western musical harmony, but much less research examines the musics of other cultures (Eerola, 2003). From the developmental standpoint, studies with infants (Trainor & Trehub, 1994) have demonstrated some knowledge of musical harmony at early ages. There has been little work, however, systematically investigating the learning of a new musical system. Thus, we take some time here to introduce the musical system used in the present experiments.

In the traditional Western musical scale, frequencies of notes are based around an octave, which is defined as a 2:1 ratio in frequency. Within the octave, the Western scale is logarithmically divided into 12 even increments. Thus, the frequencies of tones spanning an octave in the Western scale are defined as the following:

$$\text{Frequency (Hz)} = k*2^{n/12}$$

[1] This does not deny the existence of underlying physiological features that constrain possible musics. There may be both physiological processes and powerful learning mechanisms that simultaneously constrain possible musical systems.

where n is the number of steps along the chromatic scale, and k is a constant and typically equals 440Hz.

For the following experiments we used the Bohlen-Pierce scale (Walker, 2001), a microtonal tuning system based on 13 logarithmically even divisions of a *tritave*, which is a 3:1 ratio in frequency. The tones in one tritave of the Bohlen-Pierce scale are defined as:

$$\text{Frequency (Hz)} = k*3^{n/13}$$

where n is the number of steps along the tritave scale, and k is a constant which equals 220Hz.

From the above formula we defined the following pitches in one tritave of the Bohlen-Pierce scale (see Table 1):

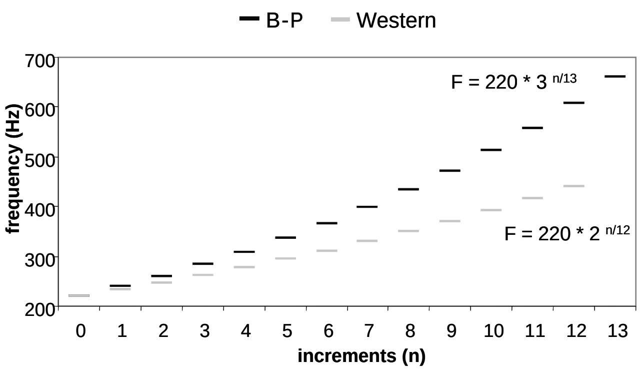

Figure 1. Frequencies along the Bohlen-Pierce scale and Western scale.

Based on this scale, it was possible to define chords using pitches with frequencies that approximately related to each other in low-number integer ratios, which are relatively consonant psychoacoustically. One "major" chord in this new system is defined as a set of three pitches with frequencies that approximate a 3:5:7 ratio (see Krumhansl, 1987, for a derivation of chords in the Bohlen-Pierce scale). We composed two sets of chord progressions, where each chord progression consisted of four chords, and each chord consisted of three pitches. Table 1 lists each of the pitches in each chord, with the frequency of each pitch given by substituting the pitch number into n in the tritave formula: Frequency (Hz) = $220 * 3^{n/13}$.

Table 1. Pitches in each chord progression.

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

These chord progressions formed the bases of the two sets of grammatical rules which we used to form the stimuli of the present study. By applying these rules in the fashion of a finite-state grammar, we composed sets of melodies from the above chord progressions. Figure 2 illustrates the chord progressions as two finite-state grammars, whereas Figure 3 provides one example of using Grammar I to generate a melody.

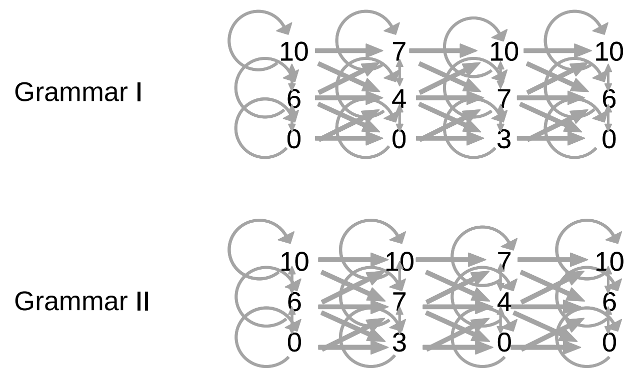

Figure 2. A finite-state grammar diagram illustrating the possible ways to compose melody from harmony.

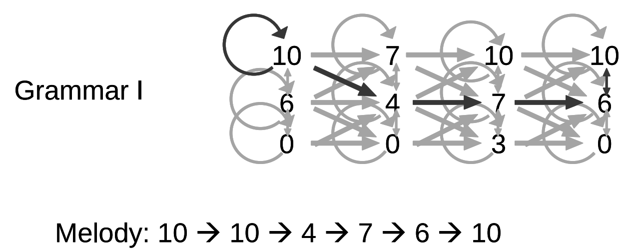

Figure 3. An illustration of applying a finite-state grammar as a set of rules to compose one melody based on finite-state Grammar I. Dark arrows illustrate the paths taken, whereas light arrows illustrate the other possible paths that are legal in the grammar. The resultant melody is shown at the bottom of the figure.

Each melody was composed according to the two grammars, with the following additional constraints:

1. Each melody ranged from four to eight notes.
2. Large intervals in one direction were followed by small intervals in the other direction (see Narmour, 1990).

Using these two finite-state grammars as compositional rules, a small subset of melodies can be legal exemplars of both Grammar I and Grammar II, but these melodies were not used in the present study. The musical grammars described here can support important investigations not only in music cognition, but also in statistical learning more generally. Materials constituting our finite-state grammar are such that the items themselves (the tones) are non-rehearsable in the sense that one cannot easily designate a verbal label to each individual tone. Unlike existing artificial grammar studies, such as those using artificial speech syllables (e.g. Gomez & Schvaneveldt, 1994), these artificial musical grammars force the cognitive system to represent items in a nonlinguistic manner, thus minimizing the possibility of rote memorizing the items of an artificial grammar through covert verbalization of its exemplars.

Having defined two sets of artificial musical grammars based on a new musical system, we now have a tool to answer many research questions. We present two experiments in which melodies composed in the musical grammars are presented repeatedly to participants, and pre- and post-exposure tests are conducted to assess the extent to which participants acquired their grammar. In particular, we ask the following questions regarding learning:

1. Can participants recognize grammatical items (melodies) they had heard?
2. Can they generalize their knowledge of the grammatical rules towards new instances of their grammar?
3. Can they form expectations for the underlying statistics of the new musical system?
4. Can they learn to form preferences for any aspects of the musical grammars?

To address questions 1 and 2, we employ two-alternative forced choice tests of recognition and generalization, where participants hear familiar and unfamiliar melodies in both grammars, and are asked to choose the more familiar melody. Question 3 is addressed using probe-tone tests before and after the exposure phase, where participants hear a melody followed by a probe tone, and rate the degree to which the probe tone fits the preceding melody. Finally, to answer question 4 we conduct a subjective rating task where participants rate their preference of each melody after hearing it.

Method

Experiment 1

Subjects 24 undergraduate students from the University of California at Berkeley participated in this study for course credit. All participants had normal hearing and more than five years of musical training. Each subject was randomly assigned to one of the two grammars.

Stimuli All auditory stimuli were generated and presented using Max/MSP (Zicarelli, 1998) and presented via headphones at a level of 70dB. 20 individual melodies (10 in each grammar) were constructed and presented in pure tones ranging from 220Hz to 660Hz, spanning one tritave in the current musical system. Each tone lasted 500ms, with envelope rise and fall times of 5ms each. The musical grammars and melodies generated from them are described in the previous section.

Procedure Experiments were run in a soundproof chamber. Each experiment was run in five phases:

1) Pre-exposure probe tone ratings: A melody was presented followed by a probe tone. The participant's task was to rate how well the probe tone fit the preceding melody, on a scale of 1 to 7 (where 1 was least fitting and 7 was best fitting).
2) Exposure: The second phase consisted of five melodies generated from the assigned grammar being played in randomized order for 25 minutes. Equal numbers of participants were exposed to the two grammars. Each tone lasted 500ms, and a 500ms silence separated any two melodies. Each melody was repeated 100 times over the course of exposure. While listening to the melodies, participants were given the option of drawing on provided paper as a distracter task.
3) Forced-choice recognition and generalization: In the third phase, participants were given a two-alternative forced-choice task. The first part of the forced-choice trials tested for participants' recognition of melodies they had heard during the exposure phase, whereas the second assessed whether they could generalize their recognition to novel melodies in the same grammar. One block of five recognition trials preceded another block of five generalization trials. In both types of trials, two melodies were played one after another, with an inter-onset interval of 5 seconds. Participants' task was to indicate which melody (the 1st or 2nd) sounded more familiar. The alternative melody (the incorrect answer) was always a melody drawn from the alternative grammar. Thus, the right answer for one group of participants was the wrong answer for the other group.
4) Post-exposure probe tone ratings: In order to measure the effect of exposure, (i.e. whether or not participants had learned anything about melodic structure) we gave them a second probe tone rating task after exposure. This task was identical to the first probe tone task, allowing direct comparison between the two sets of data.
5) Preference ratings: To assess the degree to which learning influences musical preference, the fifth and final block consisted of preference ratings: participants rated each melody in both grammars (20 melodies in all) after it was played once. Ratings were on a scale of 1 to 7, with 1 being the least preferable and 7 being the most preferable.

Experiment 2

Subjects 24 undergraduate students participated in this study. Recruitment criteria were the same as Experiment 1. None of the participants in Experiment 2 had been in Experiment 1.

Stimuli In addition to the 10 melodies for each grammar from Experiment 1, 10 new melodies were generated for each of the two grammars. All other stimulus parameters were the same as Experiment 1.

Procedure Experiment 2 was identical to Experiment 1 in procedure, except for the following modifications:

1) Pre-exposure probe tone ratings: a different melody (i.e. one not presented during the exposure) was used to obtain probe tone profiles in both pre- and post-exposure ratings.
2) Exposure: we increased the set of presented melodies from five to 15. Each melody was presented 33 times over 25 minutes of exposure.
3) Forced-choice recognition and generalization: 10 more trials were added to the recognition block, such that this phase of the experiment contained 20 trials, 15 in the recognition block and 5 in the generalization block.
4) Post-exposure probe tone ratings: this was identical in materials and procedure to the Pre-exposure probe tone ratings phase of Experiment 2.
5) Preference ratings: participants rated each of the 40 melodies (20 in each grammar) for their preference, using the same scale and procedure as Experiment 1.

Results

Experiment 1

Figure 4 shows the data from the recognition and generalization tasks, plotted as percent correct. A participant's response was scored as correct when they selected either the melody that they had heard during exposure (recognition items) or the novel melody generated with the same grammar as their exposure melodies (generalization items). When the participant selected the melody generated by the other, non-trained grammar, their answer was scored as incorrect.

As is evident in the figure, participants were significantly above chance in identifying the melodies they had heard, but were not above chance in being able to identify new instances of the same grammar as more familiar.

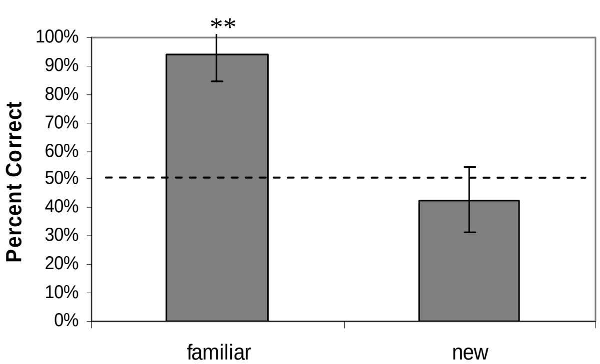

Figure 4. Two-alternative forced choice results from Experiment 1.[2]

Pre- and post-exposure probe tone ratings both revealed some sensitivity to the statistics of the melodic structure. Participants' judgments of individual probe tones were correlated with the melodic structure present in their input, and there was a trend towards an increased correlation after the exposure.

Results from the preference ratings revealed mixed effects of exposure on subjective preference. Participants exposed to one of the grammars did not significantly prefer experienced melodies to unfamiliar ones. The other participants, however, rated the melodies they had heard as significantly more preferable than melodies they had not heard. Ratings for novel melodies from the familiar grammar were not reliably different from ratings for melodies in the unfamiliar grammar (see Fig. 5), suggesting that a preference change was associated with familiarity but did not extend to other instances of the grammar.

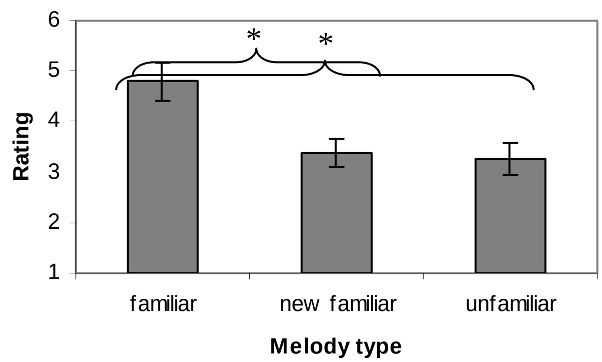

Figure 5. Preference ratings.

Experiment 2

Figures 6 and 7 show results from the recognition and generalization tasks following exposure to 15 melodies. Data are shown separately for the two groups of participants based on the presentation grammar due to a significant main effect of grammar. Forced-choice tests showed that both groups of participants recognized the presented melodies. In addition, participants exposed to Grammar I significantly generalized their familiarity to the new instances of the same grammar (see Fig. 6), whereas the group exposed to Grammar II did not generalize their familiarity to new melodies (see Fig. 7).

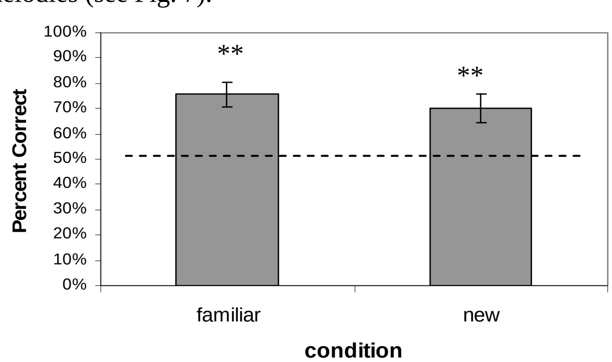

Figure 6. Grammar I forced choice results.

[2] For all figures, ** = p<0.01; * = p<0.05

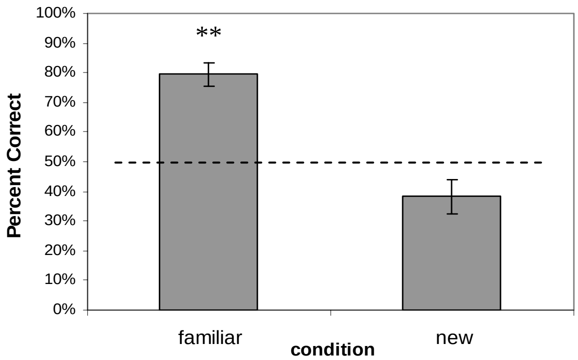

Figure 7. Grammar II forced choice results.

Probe tone tests revealed that subjects were able to pick up on the statistics of the corpus of melodies. Both the pre-exposure ratings (Fig. 8) and the post-exposure ratings (Fig. 9) were correlated significantly with the melody set, but the post-exposure ratings had significantly higher correlation. Figure 10 compares the correlations of the probe tone profiles from Figure 8 (exposure versus pre-exposure ratings) and Figure 9 (exposure versus post-exposure ratings), showing the relative statistical sensitivity reflected by pre- and post-exposure ratings. An increase in correlation is observed, suggesting that given 25 minutes of exposure to 15 exemplars, subjects became more sensitive to the underlying statistics of the new musical system.

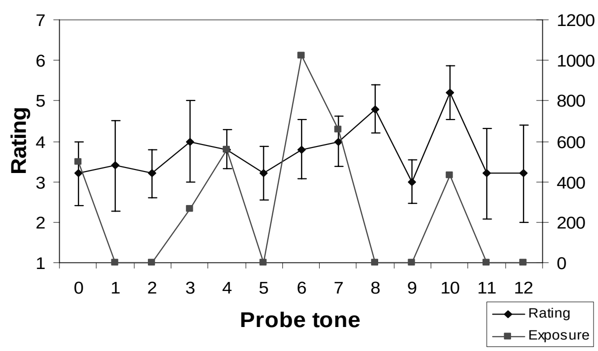

Figure 8. Pre-exposure probe tone ratings versus frequencies of occurrence of pitches during exposure.

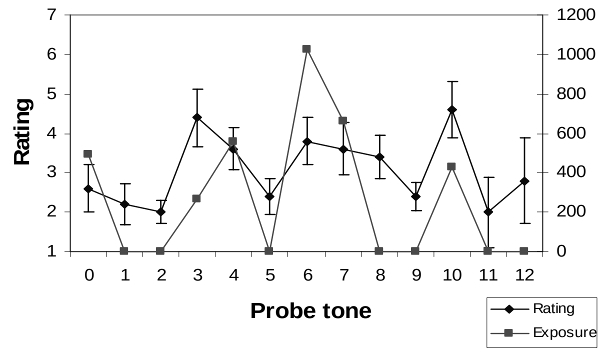

Figure 9. Post-exposure probe tone ratings versus frequencies of occurrence of pitches during exposure.

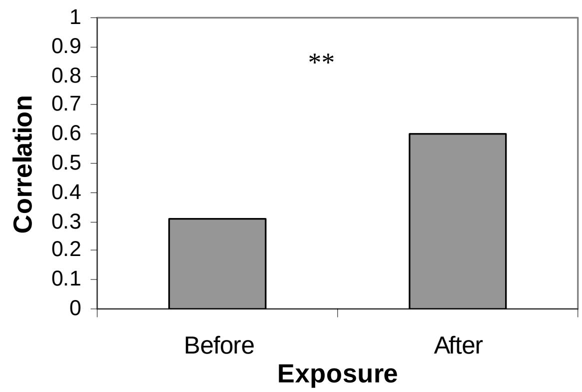

Figure 10. Correlations between profiles of probe tone ratings and frequencies of exposure, before and after exposure.

Preference ratings now did not show any differentiation between familiar, generalized, and unfamiliar melodies (see Fig. 11), suggesting that familiarity was necessary for a change in preference.

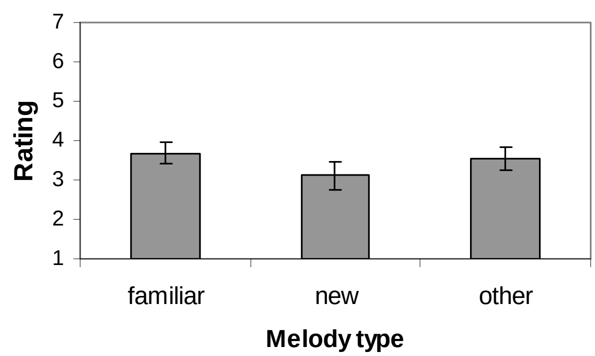

Figure 11. Preference ratings from Experiment 2.

Discussion & Conclusions

Experiment 1 showed that after participants were exposed repeatedly to five melodies for 25 minutes, they recognized and preferred melodies they had heard, but could not generalize their knowledge to new melodies composed in the same grammar.

In Experiment 2 we exposed participants to a larger number of melodies (15 instead of five) for the same overall length of time. Thus, participants heard each individual melody fewer times than participants in Experiment 1. This increase in exemplars and reduction in repetition affected participants' performance in the forced-choice tests: participants now showed not only reliable recognition, but also some evidence of generalization. Learning of the musical system was also demonstrated in the probe tone task, where the increase in correlation for post-exposure ratings demonstrated that participants became more sensitive to the underlying statistics of the grammar after exposure.

Taken together, these experiments suggest that given limited exposure to a novel musical system, humans can exhibit rapid statistical learning to develop expectations that conform to the musical grammar. Increasing the number of exemplars of the musical grammar may a) help to enhance sensitivity to the underlying statistics of a grammar; and b) aid learners in generalizing the acquired statistics to new instances of the same grammar. However, increasing the set of exemplars also seems to be detrimental towards preference ratings for familiar melodies. These results may suggest that the subjective experience of musical preference is not directly related to the increase of statistical sensitivity or to the understanding of grammatical structure; instead, preference may be a result of recognition and familiarity. We plan to conduct follow-up studies to further examine the link between familiarity and preference.

Having demonstrated that humans are able to learn new musical grammars, it is now possible to tease apart individual components of the input that constrain learning. In our ongoing research we observe reliable generalization using a much larger corpus of melodies at exposure. We have also conducted similar studies on participants with no formal musical training, with results showing that nonmusicians behave very similarly as musicians in the learning of the new musical grammars; thus we believe that these effects are not a result of formal musical education. Ongoing work also involves the use of Event-Related Potentials to study human physiological activity during the perception of this new musical system. Results from ERP studies may shed light on the neural mechanisms that enable statistical learning.

Acknowledgements

The current study was funded by a grant to D.W. from the UC Academic Senate. We thank E. Wu and P. Chen for data collection and helpful suggestions on an earlier draft of this manuscript, to C. Krumhansl for the design of the artificial grammar, and to E. Hafter for valuable advice and use of laboratory space.

References

Bharucha, J.J., Stoeckig K. 1986. Reaction time and musical expectancy: priming of chords. *Journal of Experimental Psychology: Human Perception & Performance, 12, 4,* 403-410.

Eerola, T. 2003. The Dynamics of Musical Expectancy – Cross-cultural and Statistical Approaches to Melodic Expectations. Doctoral dissertation.

Gomez, R. L., & Schvaneveldt, R. W. (1994). What Is Learned from Artificial Grammars? Transfer Tests of Simple Association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(2), 396-410.

Hunt, R.H. & Aslin, R.N. 2001. Statistical Learning in a Serial Reaction Time Task: Access to Separable Statistical Cues by Individual Learners. *Journal of Experimental Psychology: General, 130(4),* 658-680.

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

Krumhansl, C. L.; Kessler, E. J. (1983). Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys. *Psychological Review, 89(2),* 334-368.

Krumhansl, C.L. (1987). General properties of musical pitch systems: Some psychological considerations. In J. Sundberg (Ed.), *Harmony and Tonality*, Stockholm : Royal Swedish Academy of Music, publ. no. 54.

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

Meyer, L. (1956). *Emotion and Meaning in Music.* University of Chicago Press.

Narmour, E. (1990) *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. Chicago: University of Chicago Press.

Saffran, J.R., Aslin, R.N., & Newport, E.L. 1996. Statistical learning by 8-month old infants. *Science, 274,* 1926-1928.

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

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

Walker, E. (2001). The Bohlen-Pierce Scale: Continuing Research. http://www.ziaspace.com/elaine/BP/

Zicarelli, D. (1998). An extensible real-time signal processing environment for Max. *Proceedings of the International Computer Music Conference, 1998.*