2012

Are Implicit Learning Abilities in Language and Music Related?

Dana Boebinger

Abstract

One active area of psycholinguistics research seeks to determine whether language acquisition occurs through domain-specific processes or domain-general processes. In other words, are the cognitive resources dedicated to language learning specific to language, or are they part of a more general learning mechanism? One intriguing strategy for investigating this question is through the comparative use of language and music: each is a complex system in which basic auditory elements are combined in hierarchical sequences governed by abstract rules. To understand either language or music, a person must learn these syntactic rules and generalize them to new situations. This study used an artificial grammar construct to investigate whether people learn and process musical and linguistic syntax similarly, and whether there is a correlation between a person's ability to learn complex grammatical systems in multiple modalities. The effect of previous musical experience on sequence learning was also examined. Results showed virtually no correlation between the three implicit learning tasks, suggesting that there is no domain-general implicit learning ability.

Keywords: implicit learning, artificial grammar, statistical learning, language, music, individual differences

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS & SCIENCES

ARE IMPLICIT LEARNING ABILITIES IN LANGUAGE AND MUSIC RELATED?

By

DANA LYNN BOEBINGER

A Thesis submitted to the
Department of Psychology
in partial fulfillment of the requirements for graduation with
Honors in the Major

Degree Awarded:
Spring, 2012

The members of the Defense Committee approve the thesis of Dana Lynn Boebinger defended on January 31, 2012.

Dr. Michael Kaschak
Thesis Director

Dr. Nancy Rogers
Outside Committee Member

Dr. Colleen Kelley
Committee Member

Introduction

The comparison between linguistic and musical syntax is not new. Ever since Noam Chomsky's generative linguistic theory (1972), researchers have made analogies between linguistic grammar and musical structure. Fred Lerdahl and Ray Jackendoff's major contribution to the field of music cognition was their book *A Generative Theory of Tonal Music* (1983), which applied Chomsky's rules of linguistic structure to the domain of music. Lerdahl and Jackendoff believed that comparisons between language and music were largely superficial; however, research on the relationship has been continued by psychologists, musicologists, music theorists, linguists, and neuroscientists.

In order to justify a comparison, it is first necessary to present an overview of both linguistic and musical syntax. In language, phonemes are combined to make words, which are combined into phrases, which are combined into sentences, which are assembled into discourse. In a similar way, music is made up of tones, which are combined to make chords, which are combined to make chord progressions, which are assembled into entire pieces.

One principal feature of syntax is that relationships between events are not based solely on adjacency. Dependency relationships are hierarchical, and based on phrases that are combined into sentences. Like language, events in music are also hierarchically arranged. Event hierarchies in music are based on both ornamentation and tension. Certain notes or chords in a progression are structurally central, and other notes serve as elaborations of these more structurally salient units. Tension and resolution are also hierarchically related in music. These cycles of tension and resolution are associated with perceived motion and closure in music.

One of the hallmarks of linguistic syntax is the ability of the same word to take on multiple roles depending on its context (i.e. subject, object, indirect object). This property is also evident in music, though in a different way: a long and ongoing debate in music cognition centers on whether syntactic relationships in tonal music are based on cognitive relationships or psychoacoustic properties of sound. In numerous studies, evidence has been found to support each side. On one hand, particularly strong evidence for the cognitivist viewpoint is that certain functions of musical elements derive their function from the context. For example, the same chord can take on two very different functions depending on the prevailing key. The chords G-B-D and C-E-G can be a V-I progression in C major, or a I-IV progression in G major. In the first case, C major, it would sound closed and conclusive. But in G major, the same progression would lead to the phrase sounding unfinished. Multiple studies have examined the effect that this context has on syntactic processing, and shown that even musically untrained individuals are able to detect this context-dependent function (Bigand, Tillmann, Poulin, D'Adam, & Madurell, 2001).

While this supports the idea that the structures of music and language are closely related, there are obviously several fundamental differences between syntactic structure in language and music. Despite the context-dependent function of musical entities, there is no musical equivalent of grammatical categories in language, such as nouns, verbs, and adjectives. There are also fundamental differences concerning hierarchical organizations. In language, long-distance dependencies are obvious to every normal listener. However, in music, the long-distance dependencies between tension and relaxation are more subjective, and perception of these dependencies varies among listeners. Another difference between domains is the toleration of syntactic ambiguity. In language, syntax seeks to avoid ambiguity, and the goal of the cognitive system is to find one single analysis for a particular sentence. In music, syntactic ambiguity is often exploited to increase interest and complexity. For example, a chord can simultaneously function in two separate keys, and its role need never be disambiguated.

However, the differences between the specific representations of language and music are not so great that a comparison between the fundamental cognitive processes cannot be made. The useful cognitive similarities between the two domains include the presence of combinatorial organization, hierarchical and recursive structure, and functional categories that can be occupied by different entities. These more abstract similarities suggest that the human mind has basic syntactic organizational abilities, and that these abilities are not reserved solely for language. The extent to which these abilities are domain-specific or domain-general continues to be a controversial issue, and recent research has focused on cognitive and neural processes involved with the processing of musical and linguistic syntax and their degree of overlap.

Numerous studies have examined the overlap of linguistic and musical syntactic abilities—and have reached conflicting conclusions. Neuropsychology has collected many anecdotal examples of a double dissociation between musical and linguistic processing, which implies that the processing mechanisms for language and music are independent. However, neuroimaging has provided evidence that syntactic processing activates similar brain regions for both music and language. There is a large body of fMRI research that shows both music processing and language processing activate Broca’s area and other premotor regions in the brain. One of the first studies to directly compare syntactic processing of language and music was that of Patel, Gibson, et al. (1998). This study showed that out-of-key chords elicited an event-related potential (ERP) component similar to those generated by syntactic violations in language. Incongruities in both domains elicited a P600, which had previously been thought to be specific to language processing.

This study led Patel to propose his shared syntactic integration resource hypothesis (SSIRH), which postulates that although language and music have separate stored syntactic representations, there are common neural resources involved in integrating incoming elements and activating stored representations. This claim attempts to explain the dissociations in neuropsychology by stating that damage to the stored representations would result in domain-specific difficulties. One prediction resulting from this hypothesis is that if musical and linguistic syntactic integration rely on common neural resources, and neural resources are limited, there should be interference between the two. Numerous studies have found such interference (Koelsch, Gunter, Wittforth, & Sammler, 2005; Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Slevc, Rosenberg, & Patel, 2009). Additionally, this interference seems to be specific to syntax. Musical syntactic integration and linguistic semantic integration show no such interference (Slevc et al., 2009)

Statistical Learning and Artificial Grammar

Rather than measuring interference during concurrent tasks, an alternative strategy to examine the domain-specificity of linguistic syntactic processing is to look at the learning processes involved. In infants, language and music processing show similar patterns of development. Infants' perceptual abilities begin as culturally independent, but are then modified as children are exposed to their culture's language or music. The amount of overlap between learning mechanisms involved with language and music is unknown, although several studies have examined this relationship using statistical learning.

Statistical learning is a common research tool used to explore language acquisition in both infants and adults. During language learning, infants are exposed to a huge amount of linguistic input. Within this speech stream, there are measurable statistical regularities that give rise to recurring sequences. For example, given a specific phoneme, certain phonemes are more likely than others to follow. Saffran et al. (1996) made a strong case for experience-dependent learning of these regularities by showing that 8-month-old infants are able to segment words from a continuous speech stream based solely on the statistical relationships between sounds, and that this learning took place after only two minutes of exposure. Further studies (Aslin et al., 1998) established that the primary cue used by infants is the transitional probability between syllables. In this same study, infants were able to differentiate between trisyllabic nonsense words and part-words (trisyllabic sequences spanning word boundaries). These results illustrate infants' abilities to extract reasonably complex information from sensory input based solely on statistical information.

Once it was determined that infants were capable of learning complex statistical information in a speech stream, researchers sought to determine whether these abilities were confined to language. One expansion of this line of research examined infants' ability to learn sequences of musical tones. Saffran, Aslin, and Newport (1999) translated the same "language" used in their earlier experiment into "tone words," by substituting a musical note for each syllable. Just as before, the transitional probabilities between tones were the only information that listeners received. The results for both adults and infants were indistinguishable from those obtained using linguistic syllables. These findings prompted the authors to suggest that linguistic stimuli were not privileged above tones when it comes to statistical learning, and perhaps there was an overlapping learning mechanism responsible for both linguistic and musical information.

However, these experiments using statistical learning focus on the task of segmentation, a very general problem that is not present only in language. Many tasks involve pattern detection, and language acquisition is far more complicated than first-degree transitional probabilities. In order to understand how learners acquire more complex syntactic rules, researchers have used finite-state artificial grammars with more complex organizational schemes. An artificial grammar is made up of nodes connected in a specific way, and connecting the nodes in different combinations yields different “legal” sequences. There is a long history of implicit learning research that use artificial grammars, beginning with Reber’s (1967, 1969) experiments, in which participants learned to classify strings of letters as either grammatical or ungrammatical, without being able to explicitly explain the rules governing the grammar. These findings have been duplicated several times, including in other modalities, using visuospatial and visuomotor tasks.

Only a few artificial grammar studies have been conducted using musical stimuli. Jonaitis and Saffran (2009) used an artificial grammar and showed that adult listeners were able to rapidly learn to discriminate between grammatical and ungrammatical chord sequences, and that the learning process seemed to parallel that of language learning. After extended exposure to the stimuli, participants were able to use the learned statistical structures to understand the structure of novel sequences. Loui and Wessel (2009) created a microtonal harmonic artificial grammar, and created melodies as legal exemplars of this grammar. Participants were better able to extract rules and apply them to novel melodies when they were exposed to a larger set of exemplar melodies. These studies serve as examples that learners are able to extract grammatical rules from musical input.

Studies by Conway and Christiansen (2005, 2006) have compared statistical learning abilities across different sensory modalities. One of their studies used a single artificial grammar that was “translated” into an auditory, a visual, and a tactile task. Participant performance indicated that auditory information was much more easily learned. Contrary to previous studies, Conway and Christiansen found a clear difference between performance on visual and auditory tasks. They explain this by proposing that statistical learning is modality-specific: visual information is best learned in a spatial presentation, and auditory information is best learned in a temporal presentation.

The present study will examine whether a single person's artificial grammar learning abilities correlate across modalities. It has been shown that auditory information has an advantage over visual information (Conway & Christiansen, 2005), but does this advantage hold for both language and music? And if so, is this advantage the same for both music and language, or is one more privileged than the other? To explore this issue, participants will be given three implicit learning tasks: one that uses linguistic syllables, one that uses musical materials, and one that measures visuospatial learning. Participants' scores on each task will be compared to determine whether performance on a linguistic artificial grammar task correlates with performance on a musical artificial grammar task, and whether this correlates with a more domain-general visual implicit learning task. The effect of musical experience on implicit learning abilities will also be examined.

Based on previous statistical learning research, we expect that participants will successfully learn the grammatical structure in each modality, and will display some rule generalization. And according to the literature on the relationship between linguistic and musical syntactic processing, we expect that learning will be similar in these two domains. Several more recent studies show that auditory information is more amenable to statistical learning, so we also expect that participants will perform better on the musical and linguistic tasks than the visual task. What remains to be seen is whether each participant's performance correlates across domains.

Method

Participants

Sixty-four Florida State University undergraduates participated in this experiment (20 males and 44 females; *M* = 18.66 years, *SD* = 1.53 years). Participants consisted of general psychology students (n = 32) and music majors (n = 32). Music majors had an average of 9.59 years of musical experience (*SD* = 3.12). All participants reported that they had normal hearing. Psychology students participated in the study for course credit, and music majors were compensated for participation.

Materials

Linguistic stimuli. Linguistic stimuli consisted of strings of three to five nonsense syllables generated based on an artificial grammar devised by Kaschak and Saffran (2006):

(A) Sentence (S) = A-phrase (AP) + C-phrase (CP) + E-word

(B) AP = A-word + optional D-word

(C) CP = C-word + optional G-word

Each word category consisted of a set of nonsense syllables. There are four A-words (biff, hep, mib, rud), two D-words (klor, pell), four C-words (cav, lum, neb, sig), two G-words (tiz, pilk), and four E-words (jux, vot, loke, dupp).

The following sequences of word categories formed acceptable, grammatical sequences:

A-C-E

A-D-C-E

A-C-G-E

A-D-C-G-E

The linguistic stimuli used in both the training and test phases of this experiment were appropriated, with minor modification, from the materials used by Jones and Kaschak (2009).

Musical stimuli. Musical stimuli were generated following the same artificial grammar as the linguistic stimuli (Kaschak & Saffran, 2006), but substituting a specific musical trichord for each syllable.

A trichord is defined as any set of three different pitch classes. In this experiment, each trichord consisted of three pitches played concurrently. To avoid the possibility that participants heard an obvious bass or melody line, Shepard tones were used to construct each trichord. A Shepard tone is a sound created by a superposition of sine waves separated by an octave, thus eliminating the notion of a “bass” note. Eight unique trichords were generated at two transpositional levels, which amounts to sixteen total trichords. Efforts were made to use each possible pitch an equal number of times and maximize the distinctness of each trichord.

A computer program was designed to produce every possible grammatical sequence generated by the artificial grammar. A sample was then randomly selected that matched the characteristics of the linguistic stimuli. However, no sequence used in the musical sample was identical to a sequence from the linguistic sample.

For more information about the linguistic and musical stimuli, see Appendix A.

Visuospatial stimuli. Visual implicit learning was evaluated using a serial reaction time task (SRTT) that presents 3- to 8-item sequences of spatially arranged colored squares (red, green, yellow, blue), much like the popular “Simon” game. Unbeknownst to the participant, the pattern of squares follows an underlying artificial grammar.

Procedure

Each participant completed three artificial grammar tasks: one linguistic, one musical, and one visuospatial. The order of the tasks was counterbalanced between individuals. Each participant was tested individually on a PC laptop in a test carrel. All three tasks were presented using E-Prime 2.0 software. Auditory stimuli were presented via high-quality headphones.

Both the linguistic and musical task had two main parts: a training phase and a test phase. At the beginning of the training phase, participants were told that they would hear a presentation of a made-up linguistic/musical system. They were instructed to follow along with the recording and pay attention to each syllable/chord. No indication was given that the sequences followed a set of rules. During the training phase, participants were presented with a series of fifty grammatical sequences that was repeated four times. For both tasks, the training exposure lasted about 7-7.5 minutes.

The training phase was immediately followed by the test phase, which consisted of 15 pairs of grammatical and ungrammatical sequences. Ungrammatical sequences were generated by breaking one of the five core rules from the Kaschak and Saffran (2006) grammar:

(1) All sentences must have an A-phrase.

(2) In an A-phrase, A-words precede D-words; in a C-phrase, C-words precede G-words.

(3) Sentences must have an E-word.

(4) C-phrases must precede E-words.

(5) If there is a G-word, there must be a C-word.

In each test trial, the participant was required to make a forced-choice decision about which sequence in the pair sounded like it could have come from the previous presentation. Participants responded by pressing a key on a computer keyboard that corresponded to either the first or second sequence presented. The first and second sequence were correct an equal number of times. For both the musical and the linguistic task, the test phase itself took about 7.5 minutes.

The visuospatial task was structured somewhat differently. For each trial, the participant watched the presentation of the sequence, and then clicked on the colored squares in the order that they appeared during the presentation. At times, the pattern of colored squares switched from grammatical sequences to random sequences. Participants' increased reaction time during these phases is an indicator of the amount of implicit learning that has occurred. The visuospatial task took about fifteen minutes.

After completing all three tasks, participants were given a questionnaire that asked how they decided which test sequence to choose during the test phase of the linguistic and musical tasks and whether they felt like they were guessing. The questionnaire also included demographic information and details about previous musical experience. The entire experiment took about forty-five minutes.

Results

For the linguistic and musical artificial grammar tasks, scores were calculated by summing the number of times the participant correctly chose the grammatical sequence during each trial of the forced choice test, and then dividing that number by the total number of test trials. For the visuospatial task, the score was calculated by multiplying each correct response on a grammatical trial by the number of items in the sequence, and then adding those values together to create a grammatical learning score weighted by sequence length. The performance on random sequences was computed in a similar way, and then subtracted from the grammatical learning score. This yielded a learning score that reflects the degree to which each participant learned and made use of the grammatical sequences throughout the task.

Descriptive statistics for all tasks are provided in Table 1. Overall, participants showed learning effects in all tasks. On the linguistic artificial grammar task, 93.75% of participants scored above chance. For the musical artificial grammar task, 92.19% of participants scored above chance. For the visuospatial implicit learning task, 90.63% of participants performed better on the grammatical sequences than the random sequences.

Table 1
Descriptive Statistics for all Implicit Learning Tasks

Task	Dependent Measure	Mean (SD)	Median	Observed Range	Possible Range
Linguistic AG	Percent correct	.732 (.132)	.733	.400 – 1.00	0 - 1
Musical AG	Percent correct	.682 (.115)	.667	.467 – .933	0 - 1
Visuospatial SRTT	Grammatical Score – Random Score	29.188 (21.577)	30.500	-16 – 93	-120 - 120

Note: AG = Artificial Grammar, SRTT = Serial Reaction Time Task

One of the main aims of this study was to examine the correlation between performances on these three different implicit learning tasks. Pearson correlations were computed between each pair of tasks, and contrary to our expectations, no significant correlations were found. The correlation between the linguistic and musical artificial grammars was very low (r = .122, p = .335), and neither the linguistic (r = .063, p = .629) nor the musical (r = .018, p = .890) task had any correlation with the visuospatial task.

Because there were no significant correlations between the tasks at a group level, the effect of musical experience on task performance was examined. Independent samples t-tests showed no difference in the performance of musicians (M = .725, SD = .122) and non-musicians (M = .729, SD = .143) for the linguistic task, $t(62)$ = -.189, p = .851, as well as no difference in the performance of musicians (M = .696, SD = .124) and non-musicians (M = .669, SD = .105) for the musical task, $t(62) = -.941$, $p = .350$. However, non-musicians ($M = .39.16$, SD = 20.824) showed better performance than musicians ($M = 19.22$, SD = 17.531) on the visuospatial task than musicians, $t(62) = 4.143$, $p < .001$. This difference between musicians and non-musicians disappeared when the data were re-analyzed using years of musical experience as a grouping variable, rather than whether or not the participant was a music major. Thus, we have no reason to believe that this difference between musicians and non-musicians on the visuospatial task has any substantial meaning.

Although there was no significant difference between musicians and non-musicians at the task level, a 2 × 2 × 5 ANOVA, with musician/non-musician as a within-subjects variable and task and rule as between-subjects variables, showed a significant main effect of rule, Greenhouse-Geisser corrected, $F(3.521, 218.307) = 2.843$, $p = .031$. We thus broke down participants' performance on the linguistic and musical tasks by the five rules from the Kaschak and Saffran (2006) grammar, to see whether musicians and non-musicians differed in their learning of various aspects of the grammars. The five rules are as follows:

(1) All sentences must have an A-phrase.

(2) In an A-phrase, A-words precede D-words; in a C-phrase, C-words precede G-words.

(3) Sentences must have an E-word.

(4) C-phrases must precede E-words.

(5) If there is a G-word, there must be a C-word.

First, the correlation of each rule between the linguistic and musical artificial grammar tasks was computed. As Table 2 demonstrates, the only significant correlation was on Rule 5. There were no correlations between the two tasks for any of the other rules.

Table 2

Correlation of Rules Between Linguistic and Musical Tasks

Rule	Linguistic Mean (SD)	Musical Mean (SD)	Correlation Coefficient (*r*)	*p*-Value
1	.786 (.246)	.771 (.269)	.021	.866
2	.664 (.343)	.656 (.228)	.122	.334
3	.802 (.275)	.714 (.269)	.008	.949
4	.852 (.224)	.708 (.224)	.071	.573
5	.495 (.333)	.562 (.276)	-.289	.019*

Note: * = $p < .05$

To examine the rule breakdown further, the performance of music majors versus non-music majors was compared for each rule. As illustrated in Table 3, the only significant difference between the performances of musicians and non-musicians was on Rule 3 in the musical artificial grammar task. This suggests that musicians and non-musicians are processing the ends of the musical sequences in different ways.

A 2 × 2 ANOVA, conducted only on the performance on Rule 3, found a significant task by musician interaction, $F(1, 62) = 5.506$, $p = .022$, which highlights this significant difference in the performances of musicians and non-musicians on Rule 3 in the musical task.

Because of the egregious violation of sphericity, the data were analyzed using a mixed model regression to see whether a correct response was predicted by task, rule, or musical experience. For all logistic regression coefficients, see Appendix B. Results showed a significant difference between the overall performance on Rule 1 and Rule 5, $z = -1.284$, $p = .027$. There was a significant interaction between musical experience and performance on Rule 3 items, $z = -1.164$, $p = .027$. When task was added as an additional predictor of performance on Rule 3, the

Table 3
Performance of Musicians and Non-Musicians by Rule

Task	Musician Mean (SD)	Non-Musician Mean (SD)	*t*-Test	*p*-Value
Linguistic AG				
Rule 1	.833 (.207)	.740 (.277)	*t*(62) = -1.531	.131
Rule 2	.656 (.346)	.671 (.350)	*t*(62) = .179	.858
Rule 3	.760 (.296)	.844 (.254)	*t*(62) = 1.209	.231
Rule 4	.852 (.228)	.852 (.228)	*t*(62) = .000	1.00
Rule 5	.510 (.349)	.479 (.327)	*t*(62) = -.370	.713
Musical AG				
Rule 1	.760 (.284)	.781 (.262)	*t*(62) = .305	.761
Rule 2	.667 (.224)	.646 (.224)	*t*(62) = -.360	.720
Rule 3	.781 (.262)	.646 (.267)	*t*(62) = -2.047	.045*
Rule 4	.719 (.226)	.698 (.230)	*t*(62) = -.366	.716
Rule 5	.552 (.312)	.573 (.243)	*t*(62) = .298	.767

* = $p < .05$ (two-tailed, $n = 64$)
Note: AG = Artificial Grammar

significance of the three-way interaction increased dramatically, $z = 1.988$, $p = .005$. This three-way interaction can be interpreted to say that musicians performed significantly better on Rule 3 items in the music task.

Discussion

The aim of the experiment described here was to investigate whether people learn and process musical and linguistic syntax similarly, and whether there is a correlation between a person's ability to learn complex grammatical systems in multiple modalities. Results show that participants were able to learn the artificial grammar in each task and perform significantly better than chance. Additionally, although the different structures of the artificial grammars and the serial reaction time task do not easily permit a direct comparison, looking at the average score across all participants for each task, it can be seen that participants indeed performed better on the auditory tasks than the visuospatial task (Conway & Christiansen, 2005). However, this difference cannot be entirely attributed to the modality of the task, since the auditory measures were artificial grammar learning tasks and the visuospatial measure was a serial reaction time task.

Statistical results also show that, contrary to our original hypothesis, an individual's performance on a linguistic artificial grammar task does not correlate with his or her performance on a musical artificial grammar task. Furthermore, neither performance on these artificial grammar tasks correlates with performance on a more general visuospatial implicit learning task. The observed violation of sphericity between tasks indicates that participants' individual performances varied between tasks, so that while the overall performance between tasks is similar when averaged over all participants, the task performance on an individual level does not show any correlation among the tasks. This would suggest that there is no single implicit learning ability that holds consistently across modalities.

The absence of correlation between different implicit learning tasks is not a new finding. Several studies have found that different implicit learning paradigms are only weakly correlated with each other (Pretz, Totz, Kaufman, 2010; Salthouse et al., 1999; Gebauer & Mackintosh, 2007). Gebauer and Mackintosh (2007) used three implicit learning experimental paradigms: artificial grammar, serial reaction time task, and what they called a "process control" task. They found no correlation between the three tasks in their experiment. They did, however, find a weak correlation (r = .27, p < .05) between the first and second presentations of their artificial grammar task. Although this correlation is stronger than any observed in our experiment, the findings of Gebauer and Mackintosh (2007) are relevant to this study.

Gebauer and Mackintosh (2007) go on to claim that perhaps the different paradigms for assessing implicit learning are, in fact, measuring different constructs. They propose that the temporal stability of performance on implicit learning tasks is low, so that participants' performances are inconsistent over time. Gebauer and Mackintosh (2007) also suggest that the low correlation between the two presentations of their artificial grammar could be due to a negative transfer between the two grammars, since they had different rules. However, that would not explain the absence of correlation observed in this experiment, since the linguistic grammar and the musical grammar had the exact same structure and rules. Furthermore, the order of the three tasks in this experiment was counterbalanced to cancel out any negative interference.

As Conway & Christiansen (2005) illustrated, there are modality constraints on statistical learning such that auditory information seems to be more suited for sequential implicit learning. These modality constraints tend to imply that statistical learning is modular and processes information from the various senses in different ways. If this is true, perhaps there are further subsystems of implicit learning for different kinds of information within a modality – like speech and music. This could account for the lack of correlation between the linguistic artificial grammar task and the musical artificial grammar task. Other findings from Conway and Christiansen (2006) show that statistical learning seems to be more stimulus-specific than is commonly believed. Rather than learning the deep structure of artificial grammars, perhaps participants are instead learning the surface-level perceptual features of the stimuli. This could help explain the results presented here, because the surface-level characteristics of the stimuli in this experiment differed between the three tasks. This would have led the participants to focus on different aspects of the stimuli, and their learning would have been qualitatively different between tasks.

It could be that implicit learning is not a single ability, contrary to what Reber would say, but rather a collection of unrelated skills that depend on the task and modality. After all, artificial grammar and serial reaction time tasks make different demands of the participants. Seger (1994) points out that the various implicit learning paradigms differ in the amount of implicit and explicit processes that are needed to perform well on the task. The tasks used in this experiment also differed in the type of learning that was necessary. The visuospatial serial reaction time task involved motor learning, and the other two were auditory and didn't require motor pattern learning. This difference and the corresponding lack of correlation between the two types of tasks is perhaps explained by the theory that there are two types of implicit learning – motor-based and judgment-based (Seger, 1994) – which matches up very well with the distinction between the visuospatial serial reaction time task and the two artificial grammar tasks. Also, the visuospatial task didn't have a separate test phase, so participants were never explicitly aware that there was a pattern nor that their knowledge of that pattern was being tested. The linguistic and musical tasks, however, had distinct training and test phases and might have required more explicit processing during the training phase and recall during the test phase.

Within the auditory modality, speech and music have fundamentally different processing demands. Music involves more spectral information, and language involves more temporal information. Perhaps participants simply used different strategies during the two tasks. One explanation of statistical learning holds that as participants become more familiar with the artificial grammar, they are able to parse the sequences into more diagnostic units (Perruchet & Vinter, 2002). Perhaps the inherent differences between language and music caused participants to form different chunks. For example, the musicians focused more on the endings of the musical sequences but not the linguistic sequences.

There is also a debate over the role that individual differences play in implicit learning. Traditionally, it has been assumed that implicit learning is largely invariant across the population (Reber, 1993). However, recent studies have found systematic variation in implicit learning abilities (Misyak & Christiansen, 2011; Kaufman et al., 2010). These studies differ in which abilities seem to correlate with implicit learning. Misyak and Christiansen (2011) found that implicit learning correlates with verbal working memory and language comprehension, while Kaufman et al. (2010) found correlations between implicit learning and verbal analogical reasoning, processing speed, performance on foreign language exams, and the personality traits of intuition, openness to experience, and impulsivity. Despite a disagreement about what skills correlate with implicit learning, it is largely agreed that implicit learning shows no correlation, or perhaps a very weak correlation, with intelligence and IQ score (Misyak & Christiansen, 2011; Gebauer & Mackintosh, 2007). Thus, our lack of observed correlation could be due to the differences in verbal working memory, processing speed, or other personality traits of our participants.

Some participants were inevitably more attentive than others, and this could have also had a big impact on how well they were able to learn the grammars and sequences. Studies have shown that it seems to be necessary for the participant to selectively attend to the stimuli for implicit learning to occur, but once they are attending, learning occurs automatically (Jiang & Chun, 2001; Kaufman et al., 2010). Although our participants were given explicit instructions to pay attention to the sequences, there is always the chance that their attention wandered during the training phase. If the degree of mind wandering differed between the three tasks, that would explain the lack of implicit learning that took place, and the lack of a correlation between the three tasks.

Another important factor to consider when accounting for the lack of correlation between tasks in this experiment is the role of experience. Interestingly, musicians were more sensitive to the endings of sequences during the music task. This was not true of musicians for the other tasks, or for non-musicians in the music task. In Western music, the ending of phrases provides important information about tonality, closure, and structure. Due to their extensive exposure to Western music, musicians might have, either explicitly or implicitly, applied their Western musical listening approach to the strange-sounding sequences in the musical artificial grammar task. Non-musicians, however, did not have this tendency. The difference in the listening strategy of musicians and non-musicians seems to be a function of musicians' increased experience with listening to musical phrases.

In conclusion, this within-subjects experiment shows that implicit learning does not seem to correlate across domains. This finding is consistent with studies that show individual differences in implicit learning and a lack of correlation between performances on different implicit learning paradigms. However, these findings seem to contradict the literature that claims that syntax in language is processed in much the same way as structure in music. Although this study is narrow in its scope and limited by its correlational nature, it adds to the growing evidence that implicit learning is not one single ability, but rather a modular system that differs by modality and task.

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321-324.

Bigand, E., Tillman, B., Poulin, B., D'Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81*, B11-B20.

Christiansen, M.H., Kelly, M.L., Shillcock, R.C., Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition, 116*, 382-393.

Conway, C. M., Christiansen M.H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17*(10), 905-912.

Conway, C.M., Christiansen, M.H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 24-39.

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

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition, 37*(1), 1-9.

Gebauer, G.F., Mackintosh, N. (2007). Psychometric Intelligence Dissociates Implicit and Explicit Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 34-54.

Jiang, Y., Chun, M.M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology, 54A*(4), 1105-1124.

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

Jones, J.L., & Kaschak, M.P. (2009). Do Idiomatic Constructions Always Aid Language Learning? *Language Learning and Development, 5*, 1-25.

Kaschak, M.P., & Saffran, J.R. (2006). Idiomatic syntactic constructions and language learning. *Cognitive Science, 30,* 43-63.

Kaufman, S.B., DeYoung, C.G., Gray, J.R., Jimenez, L., Brown, J., Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*, 321-340.

Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception: An Interdisciplinary Journal, 2*(2), 131-165.

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

Kuhn, G., & Dienes, Z. (2005). Implicit Learning of Nonlocal Musical Rules: Implicitly Learning More Than Chunks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1417-1432.

Lerdahl, F., Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

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

Louis, P., Wessel, D. (2006). Acquiring new musical grammars - A statistical learning approach. *9th International Conference on Music Perception and Cognition,* Alma Mater Studiorum University of Bologna.

Misyak, J.B., Christiansen, M.H. (2011). Statistical Learning and Language: An Individual Differences Study. *Language Learning,* 1-30.

Misyak, J.B., Christiansen, M.H., Tomblin, J.B. (2010). Sequential Expectations: The Role of Prediction-Based Learning in Language. *Topics in Cognitive Science, 2,* 138-153.

Patel, A. D. (1998). Syntactic processing in language and music: Different cognitive operations, similar neural resources? *Music Perception: An Interdisciplinary Journal, 16*(1), 27-42.

Patel, A. D. (2008). *Music, language, and the brain*. New York, New York: Oxford University Press, Inc.

Perruchet, P., Vinter, A. (2002). The formation of structurally relevant units in artificial grammar learning. *The Quarterly Journal of Experimental Psychology, 55A*(2), 485-503.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855-863.

Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious.* New York: Oxford University Press.

Reber, A.S., Allen, R. (1978). Analogy and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition, 6*, 189-221.

Saffran, J. R. (2003). Musical learning and language development. *Annals of the New York Academy of Sciences, 999*, 397-401.

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

Seger, C.A. (1994) Implicit learning. *Psychological Bulletin, 115,* 163-196.

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

Acknowledgements

This undergraduate honors thesis is the culmination of my time at Florida State University, and I want to thank the myriad of people who allowed me to forge my own path and encouraged all of my efforts along the way. First of all, I want to thank Dr. Michael Kaschak for trusting me enough to give me a large amount of freedom when it came to choosing my thesis topic and designing the experiment. I also want to thank him for constantly being willing to help when that freedom created problems for me. Dr. Nancy Rogers provided helpful advice about the musical aspects of this study, as well as support for my academic trajectory, and for that I thank her. Dr. Colleen Kelley has been a great teacher, and pushed me to work hard and achieve. I also want to thank Dr. Josh Rodefer for offering lighthearted advice about both scientific research and life in general that kept me sane throughout these stressful times. I want to thank Dr. Clifford Madsen, who has been a fabulous mentor during my tenure at FSU. His optimism about my future and his willingness to bend the rules to allow me to pursue my goals has been invaluable to me over the past few years.

I must also thank my boyfriend, Michael, for both designing the computer programs that were used in this study, as well as providing a listening ear throughout this process. And lastly, I would like to thank my parents, who are my biggest cheerleaders and who would be proud of me no matter what this document contained.

Appendix A

A.1 *Linguistic Sequences*

A.1.1 *Training Sequences*

mib pell lum dupp
hep lum vot
rud pell sig pilk dupp
biff klor neb jux
mib sig tiz vot
hep pell cav pilk vot
mib lum dupp
rud klor lum loke
hep neb vot
mib cav dupp
rud pell sig vot
biff neb pilk jux
rud klor cav vot
biff klor cav tiz dupp
mib neb loke
rud lump ilk neb dupp
biff pell sig vot
biff sig vot
hep cav jux
mib klor cav dupp
hep pell sig jux
hep klor neb pilk loke
biff lum tiz dupp
hep klor neb jux
biff sig loke
mib pell neb loke
hep cav pilk jux
mib lum tiz loke
rud sig tiz loke
hep neb tiz dupp
mib sig pilk jux
rud lum jux
mib sig dupp
rud cav vot
mib pell neb tiz jux
hep klor lum vot
biff sig tiz dupp
hep sig loke
mib cav tiz vot
biff pell sig tiz loke
mib klor lump ilk jux
rud lum loke
rud neb jux
biff lum jux
hep sig dupp
biff pell lum vot
rud neb tiz vot
he plum pilk loke
biff neb tiz loke
rud klor sig jux

A.1.2 *Test Sequences*

Correct	Error
1. rud cav pilk loke	cav pilk loke
2. biff neb pilk vot	biff neb pilk
3. biff sig pilk dupp	biff sig pilk
4. hep klor sig pilk jux	hep klor pilk jux
5. biff pell neb dupp	biff pell dupp neb
6. mib pell lum pilk dupp	mib pell pilk dupp
7. rud pell lum tiz loke	rud pell tiz loke
8. hep pell lum jux	hep pell jux lum
9. mib sig vot	mib vot sig
10. hep klor cav dupp	hep klor cav

11. rud pell cav tiz jux	rud tiz cav pell jux
12. biff klor lum tiz loke	biff tiz lum klor loke
13. biff cav dupp	cav dupp
14. hep sig tiz dupp	sig tiz dupp
15. rud neb vot	rud vot neb

A.2 *Trichords*

All trichords are named by pitch classes (C = 0, C# = 1, D = 2, etc.). Because Shepard tones were used to create each trichord, the order of the pitch classes is irrelevant. Efforts were made to choose trichords that did not closely resemble common chords in Western music (like major/minor triads and dominant 7th chords). Efforts were also made to use each possible pitch an equal number of times.

Transpositional Set #1:

8 – 9 – 10	(5)
2 – 1 – 10	(6)
0 – 1 – 6	(1)
3 – 5 – 7	(7)
6 – 4 – 11	(3)
11 – 2 – 5	(4)
4 – 0 – 8	(2)
7 – 6 – 3	(8)

Transpositional Set #2:

11 – 0 – 1	(14)
5 – 4 – 1	(16)
3 – 4 – 9	(15)
6 – 8 – 10	(10)
9 – 7 – 2	(12)
2 – 5 – 8	(13)
7 – 3 – 11	(11)
10 – 9 – 6	(9)

Each trichord was given an ID number in order to create the sequences. The number in the parentheses following each trichord is its ID number.

A.3 *Musical Sequences*

A.3.1 *Training Sequences*

For each sequence, the trichords are indicated by their ID number (see A.2 above). In the Kaschak and Saffran (2006) artificial grammar, A = trichords 1-4, D = trichords 5-6, C = trichords 7-10, G = trichords 11-12, and E = trichords 13-16.

1 – 9 – 12 – 15
3 – 10 – 15
4 – 8 – 16
4 – 10 – 14
3 – 6 – 9 – 14
3 – 9 – 13
2 – 5 – 8 – 16
1 – 6 – 8 – 14
3 – 8 – 11 – 14
1 – 9 – 12 – 16
1 – 5 – 10 – 16
3 – 6 – 8 – 15
3 – 6 – 9 – 12 – 16
1 – 7 – 16
3 – 9 – 11 – 15
1 – 9 – 15
1 – 9 – 13
4 – 9 – 12 – 15
3 – 5 – 9 – 12 – 14
2 – 6 – 7 – 12 – 15
4 – 5 – 10 – 14
4 – 6 – 7 – 12 – 15
2 – 10 – 14
3 – 6 – 7 – 16
2 – 9 – 13
4 – 6 – 10 – 14
2 – 5 – 10 – 12 – 16
3 – 6 – 9 – 11 – 16
4 – 7 – 13
2 – 7 – 11 – 16
1 – 5 – 10 – 11 – 15
1 – 5 – 9 – 16
3 – 9 – 11 – 14
3 – 10 – 16
4 – 10 – 11 – 15
4 – 6 – 8 – 11 – 14
3 – 7 – 12 – 13
2 – 9 – 11 – 14
3 – 9 – 16
3 – 10 – 13
2 – 8 – 12 – 14
1 – 10 – 13
2 – 5 – 8 – 11 – 15
4 – 6 – 8 – 13
2 – 9 – 13
2 – 6 – 7 – 12 – 15
1 – 5 – 7 – 16
2 – 6 – 9 – 15
3 – 6 – 10 – 13
1 – 6 – 9 – 11 – 13

A.3.2 *Test Sequences*

Correct:	Error:
1. 3 – 10 – 13	3 – 13 – 10
2. 1 – 6 – 9 – 15	1 – 6 – 9
3. 1 – 8 – 11 – 15	1 – 11 – 15
4. 4 – 6 – 7 – 12 – 15	4 – 6 – 12 – 15
5. 4 – 7 – 15	7 – 15
6. 2 – 6 – 7 – 12 – 16	2 – 12 – 7 – 6 – 16
7. 4 – 6 – 9 – 13	6 – 9 – 13
8. 3 – 9 – 11 – 13	3 – 9 – 11
9. 1 – 6 – 10 – 12 – 13	1 – 12 – 10 – 6 – 13
10. 2 – 7 – 11 – 13	2 – 11 – 7 – 13
11. 2 – 5 – 10 – 12 – 14	5 – 10 – 12 – 14
12. 2 – 9 – 15	2 – 15 – 9
13. 3 – 8 – 12 – 14	3 – 8 – 12
14. 4 – 7 – 12 – 16	4 – 7 – 16 – 12
15. 3 – 5 – 8 – 11 – 15	3 – 5 – 11 – 15

Appendix B

Table B.1

Logistic Regression Coefficients

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.1936	0.4238	2.816	0.00486 **
Musician	0.6137	0.3777	1.625	0.10415
Task	0.1584	0.5972	0.265	0.79081
Rule.f1	-0.4477	0.6500	-0.689	0.49101
Rule.f2	0.5665	0.6118	0.926	0.35447
Rule.f3	0.6967	0.5756	1.210	0.22614
Rule.f4	-1.2839	0.5796	-2.215	0.02675 *
Musician:Task	-0.7381	0.5097	-1.448	0.14764
Musician:Rule.f1	-0.6868	0.5296	-1.297	0.19467
Musician:Rule.f2	-1.1641	0.5265	-2.211	0.02704 *
Musician:Rule.f3	-0.6149	0.5171	-1.189	0.23437
Musician:Rule.f4	-0.4810	0.4739	-1.015	0.31007
Task:Rule.f1	-0.2830	0.8735	-0.324	0.74596
Task:Rule.f2	-1.2885	0.8458	-1.523	0.12765
Task:Rule.f3	-1.0337	0.8277	-1.249	0.21170
Task:Rule.f4	0.2662	0.8252	0.323	0.74700
Musician:Task:Rule.f1	0.9062	0.7037	1.288	0.19785
Musician:Task:Rule.f2	1.9883	0.7126	2.790	0.00526 **
Musician:Task:Rule.f3	0.8530	0.7077	1.205	0.22807
Musician:Task:Rule.f4	0.5030	0.6680	0.753	0.45145

*** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$
Rule.f1 = Rule 2
Rule.f2 = Rule 3
Rule.f3 = Rule 4
Rule.f4 = Rule 5

Note: All estimates are made in relation to Rule 1. Positive estimates indicate that the log odds of musicians obtaining a correct answer are greater than the odds for non-musicians. Negative estimates indicate that the log odds of non-musicians obtaining a correct answer are greater than the odds for musicians.