# Tacit tonality: Implicit learning of context-free harmonic structure

Martin Rohrmeier,*[1] Ian Cross*[2]

*Centre for Music & Science, Faculty of Music, University of Cambridge, United Kingdom*
[1]`mrohrmeier@cantab.net`, [2]`ic108@cam.ac.uk`

## ABSTRACT

Musical knowledge, like native language knowledge, is largely implicit, being represented without awareness of its complex structures and incidentally acquired through interaction with a large number of samples. Two experiments explore implicit learning of hierarchical harmonic structures of different complexity employing an artificial grammar learning paradigm. The experiments consisted of an incidental learning phase using a distraction task, and a testing phase employing the process dissociation procedure paradigm (Jacoby, 1991). Participants performed significantly above chance and recognised adjacent and long-distance dependencies in both experiments. Confidence ratings and inclusion/exclusion response patterns suggest that both implicit structure knowledge and explicit judgment knowledge are in operation. Participants recognised stimuli with deep structures that appeared in the learning phase better than new structures for the more complex grammar, whereas there was no such difference for the simpler grammar. They performed significantly better for the less complex grammar which indicates that grammatical complexity affects learnability and recognition performance. The results conform to other experimental findings that musicians and nonmusicians are able to perceive long-distance dependencies and embedded structures in tonal harmony.

## I. BACKGROUND

Humans possess a remarkable capacity to acquire knowledge about their environment whilst interacting with it in an unsupervised manner. We learn the subtleties of keeping balance on a bicycle, or of driving a car through complex traffic, and we acquire fine-grained linguistic distinctions or musical features (Shanks, 2005; Tillmann, 2005). Musical knowledge, like native language knowledge, is found in both musicians and nonmusicians, and is largely implicit (Bigand & Poulin-Charronat, 2006); it is mentally represented without awareness of its complex rules, being acquired through interaction with a larger number of samples. In this context, Thompson (2009: ch.4) proposes a process of musical acquisition, within which implicit learning may be argued to constitute a central process in musical enculturation.

Some research has shown that humans *possess* implicit knowledge about Western musical structure. For instance, priming studies have demonstrated participants' implicit musical, and, in particular, harmonic, knowledge (Tekman & Bharucha, 1992; Tillmann, 2005, gives an overview) and Deliège et al. (1996) found that nonmusicians as well musicians applied knowledge about harmony to a large extent. Schellenberg et al. (2005) demonstrated that children exhibit implicit knowledge of harmony. However, despite these findings, there is comparatively little research exploring the process of the implicit *learning* of musical structures, particularly with respect to the types of structures that can be acquired. Some studies have explored implicit learning of melody (Kuhn & Dienes, 2005, 2006; Dienes & Longuet-Higgins, 2004; Loui et al., 2008; Rohrmeier et al., submitted). However, little work has been done regarding the implicit learning of other basic musical structures, such as harmonic structure. This study aims to address this lacuna within the literature.

### A. Harmonic structure

Formalising Schenker's theory (1935), Lerdahl & Jackendoff (1983) proposed that Western music is organised by recursive hierarchical dependency relationships between musical elements in terms of time-span reduction and prolongation structure. There is a background of theoretical evidence that tonal harmony is organised in a comparable, hierarchical way. Early attempts by Kostka and Payne (1984: ch. 13) or Baroni et al. (1982) suggest that harmony is organised in hierarchical layers. Current theoretical approaches (Steedman, 1984, 1996; Rohrmeier, 2007; Giblin, 2008; Tojo et al., 2006; Pesetsky, 2007; de Haas & Rohrmeier, submitted) suggest that the structure of harmony exceeds the simplicity of a straightforward chord transition table (or finite-state grammar complexity, Chomsky, 1956), like Piston's table of root progressions (Piston, 1948), and may be modelled by hierarchical, context-free or phrase-structure grammars (Chomsky, 1956). Figure 1 illustrates one example of the hierarchical dependencies in a short harmony sequence. Further details and features of the hierarchical organisation of harmony are discussed below. This theoretical background motivated our experiment, investigating the learning of an artificial harmonic structure of a comparable hierarchical complexity.

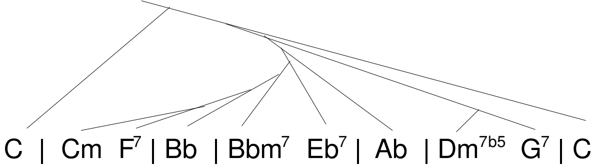

**Figure 1. Hierarchical harmonic structure in terms of time-span reduction for the first 7 bars of the Jazz standard "Afternoon in Paris" (see score in Figure 4)**

### B. Hierarchical structure and long-distance dependencies

From a different perspective, these features of hierarchical structure are linked with current advances and issues in cognitive science: the question about the learnability of recursive, hierarchical structures constitutes one of the current core topics in cognitive sciences. Hierarchy and recursion in various forms of human communication and planned action have been argued to be unique to human cognition (Fitch et al, 2005; Jackendoff, 2007) and, furthermore, Chomsky (1995) and Pinker & Jackendoff (2005) situated recursion at the heart of the human faculty of language; these debates were brought into the context of music by Jackendoff & Lerdahl (2003). Jackendoff (2007, 2009) situates the hierarchical organisation of language and music within a broader human capacity of recursion, a position that is similarly argued by Steedman

(2002).

In this context, the question of how humans form and acquire complex embedded and hierarchical structures constitutes a core question in the area. The main features that such structures embody are nested hierarchy, long-distance dependencies, and the recursive rules that create the structure. Context-free grammars, or phrase-structure grammars, constitute the simplest form of grammars to embody these features in the Chomsky hierarchy (Chomsky, 1956; Chomsky & Schützenberger, 1963).

The current empirical evidence about the learning and perception of these structures is ambiguous, and, as a result, discussion in this area is ongoing: Whereas Fitch & Hauser (2004) claim to have found evidence for learning of simple centre-embedded context-free or phrase-structure grammars in two species, Perruchet & Rey (2005), found contrasting evidence. Also, Hochmann et al. (2008) found that participants were not able to acquire specific features of the same grammar used by Fitch and Hauser. A number of other studies have investigated the learning of context-free structures in linguistic contexts: Saffran (2001; 2002) and Friederici et al. (2006), and Rohrmeier et al. (in prep.) studied human learning of phrase-structure grammars in pseudo-languages based on simple artificial words and found some evidence in favour of participants' ability to acquire the grammatical structures.

The study of musical structure is related to the issues outlined above: tonal music has been argued to embody a complex, multi-level structure that encompasses the above features of hierarchy, long-distance dependencies and recursion (Lerdahl & Jackendoff, 1983; Steedman, 1996). Patel (2003, 2007) argues that the hierarchical organisation of elements constitutes one of the structural relationships between (Western) music and language which may be linked to shared neural resources. Cook (1987), however, found some experimental evidence against participants' perception of large-scale key dependencies in music, although this study was strongly criticised by Gjerdingen (1999). Woolhouse et al. (submitted) found evidence supporting the notion that musicians and nonmusicians can perceive non-adjacent key relationships in modulating harmonic sequences. Creel et al. (2004) investigated learning of non-adjacent dependencies in tone sequences and only found learning when the long-distant dependent structures where separated from the surrounding structures in terms of auditory streaming integration.

Within these ongoing debates in the cognitive sciences and music perception, the question of the extent to which humans acquire implicit knowledge about novel, complex hierarchical harmonic structures is important. This background provided the motivation to perform a study that investigated aspects of the learning of hierarchical harmonic structures (chord sequences). Considering the debate referred above, and the fact that harmony is argued to be hierarchically organised, we decided to employ *artificial* hierarchical harmonic structures based on a novel Chomskian phrase-structure grammar. This was done in order to be able to relate the methodological framework, as well as the results, to the cognitive debates discussed above. The study employed grammars which featured aspects of hierarchical structure, centre-embedding and long-distance dependencies.

### C. Artificial grammar learning paradigm

In order to investigate the implicit learning of harmonic structure, the present research, like other cognitive studies referred to above, tied in with the well-established artificial grammar learning paradigm in psychology (Cleeremans et al., 1998; Pothos, 2007). Founded by Reber (1967), artificial grammar learning has been explored in a large variety of structures and domains, e.g. in terms of letter sequences, syllables, tones (Altmann et al., 1995), timbres (Tillmann & McAdams, 2004), or, melody (Loui et al., 2008; Rohrmeier et al., submitted). Whereas most research in artificial grammar learning has mainly focussed on simpler finite-state grammars (Chomsky, 1956), which are of a lesser complexity than the structures discussed above (Chomsky & Schützenberger, 1963), this study expanded the framework by using a formal, context-free grammar to model sequences with recursive hierarchical features.

## II. MOTIVATION

The aim of this study was twofold. Based on the scientific context outlined above, the first and main objective of this experiment was to investigate whether people could learn a novel, complex new harmonic structure that featured some of the aspects of tonal harmony, such as nested hierarchical structure, long-distance dependencies and recursion. By virtue of the structural complexity involved, the second aim of this study was to contribute more generally to the current debate in cognitive science referred to above.

According to the first aim, a generative rule system was used that produced harmonic sequences structurally commensurate with those of tonal music, but which were distinct in a number of ways. It is impossible – as in melodic learning – to avoid biases from participants' previously acquired knowledge about Western music. On the other hand, people learn to distinguish specific stylistic differences and features within Western music, such as the substantial differences between the harmony of Debussy, Chopin, Bach, Palestrina, Hindemith, Shostakovic, Prokofieff, or Ligeti (Kostka & Payne, 1984, Piston, 1948, Gauldin, 1997) and may therefore be expected to be able to acquire characteristics of a new harmonic system. This ability to learn the fine grained differences in the harmony of different composers supports learning-based accounts of music cognition (as opposed, for example, to nativist accounts). The idea that there are various features of harmony that are contingent and learned seems self-evident, even though some aspects of harmonic features may stem from psycho-acoustical bases (Parncutt, 1989).

With regard to the ability to acquire variable stylistic differences using the same musical elements, we decided to use materials that employed largely familiar chords in a new way. This avoided participants having to keeping track of and perceptually categorise a large number of new elements.

One of the main methodological challenges of this study concerned the problem of how the hierarchical dependencies that are typical in harmonic and melodic structure in tonal music may possibly be modelled in an abstract way for an experimental study. An analytical example (see Figures 2 & 3) shows the details and features of harmonic structure that this study focused on.

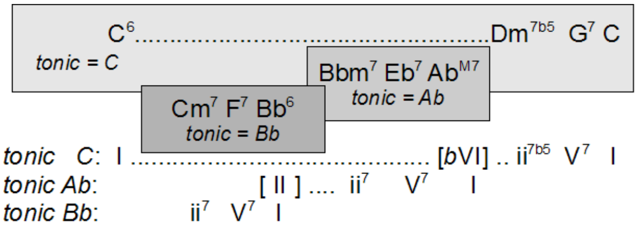

**Figure 2. Harmonic structure of "Afternoon in Paris"**

The harmonic structure of the first eight bars of John Lewis's Jazz standard "Afternoon in Paris" is complex: It features two different keys or local tonicisations of Bb major, Ab major embedded in an overarching key of C major. This creates a long-distance dependency between the first occurrence of C major and its occurrence at the end of the phrase.

The inserted Ab major segment is linked to the dominant G of the C major fragment in as much as it tonicises the flat supertonic of G, which serves as a Phrygian preparation of the dominant. The segment in Bb major is linked to the segment in Ab major by virtue of the fact that the predominant Bbm$^7$ (scale degree ii of a II-V-I sequence in Ab) is reached by its own prior tonicisation Bb, with employs an independent II-V-I sequence. The scale degree ii (Cm) of this example is reached as a minor alteration of the initial C major tonic chord of the piece. However, since the Bb sequence could similarly possibly feature a IV-V-I sequence (i.e. the chord sequence Eb-F-Bb, though less stylistically typical for this style of Jazz), the link from C to Cm only constitutes a smooth surface connection, although not a deep structural connection as there is no comparably straight-forward relationship between C and Eb for the alternative sequence. This legitimises the dependency relationships between the chords and keys as depicted in Figure 2. Without this formalisation, the deep, underlying structural similarity of the three instances of ii-V-I sequences in the example would be omitted In addition, it is important to note that the organisation of the elements within the prototypical ii-V-I sequence is hierarchical: the chord V is dependent on I, and the chord on ii is dependent on V. This can be illustrated by the fact that the sequence ii–I omitting the V chord is less regular and rarely found in Jazz harmony, whereas the sequence V – I, omitting the ii chord is a frequent alternative.

Hence the harmonic structure above embodies a similar embedded structure as to the English sentence "The book [,which contained Shakespeare's sonnet, [which I wanted to read,]] was not on the shelf." (the dependency structure is similar but reversed: dependent relative clauses bind to the left and not to the right). Such complex structural dependencies are found in Jazz standards (Burbat, 1988; de Haas & Rohrmeier, submitted).

In summary, this example features several prototypical features of tonal harmony:

- There are three different functional categories of harmonies which feature different chord instances, such as tonic, dominant (e.g. V or *b*II) and subdominant (eg. IV, II, ii$^0$, VI).
- There is a hierarchical embedding dependency relationship between key structures or local modulations. These hierarchical relationships are recursive (Tomalin, 2007), as has been stated for tonal music (Lerdahl & Jackendoff, 1983; Baroni, 1983; Steedman, 1984; Rohrmeier, 2007).
- There are hierarchical dependencies between structural elements.

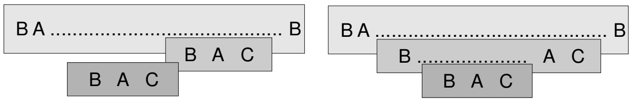

**Figure 4. Representation of the embedded structures of grammar 1 (left) and 2 (right)**

These aspects of harmony have been theoretically formalised and described for harmony in Jazz (Steedman, 1996; Chemiller, 2004; Burbat, 1988; Levine, 1996) and classical music (Kostka & Payne, 1984; Rohrmeier, 2007).

All of these features constitute a structure that cannot be sufficiently expressed or modelled with a transition table like Piston's table (Piston, 1948), or a finite-state grammar structure (the full argument relating to this point has been omitted due to space limitations, but is given in Rohrmeier, 2007). This observation underpinned the rationale of this study, as outlined above, and motivated us to use a structure of a similar kind of complexity.

A study by Woolhouse et al. (submitted) used a setup which modelled embedded key relationships of the form *key 1 – key 2 – key 1,* similar to the structures in Figure 1. The study found that participants could recognise long-distance key dependencies and that the recognition decreased as a function of the time interval between both instances of *key 1.*

### A. Design of the grammar

With regard to the discussion above and the methodology used by Woolhouse et al., the study here chose to use materials of the same complexity and with similar core features as exhibited by tonal harmony. Therefore we employed a grammar that was *structurally similar* to the outlined features of tonal harmony (and with regards to the nature of the rules employed), yet *distinctly different* on the surface and implementation level. It featured three classes of chords, and rules, that described their mutual dependencies as well as sequence centre-embedding.

Therefore, the grammar featured 3 classes of elements, A, B and C the order of which constituted the deep structure of the sequences. The deep structure was created so that it featured
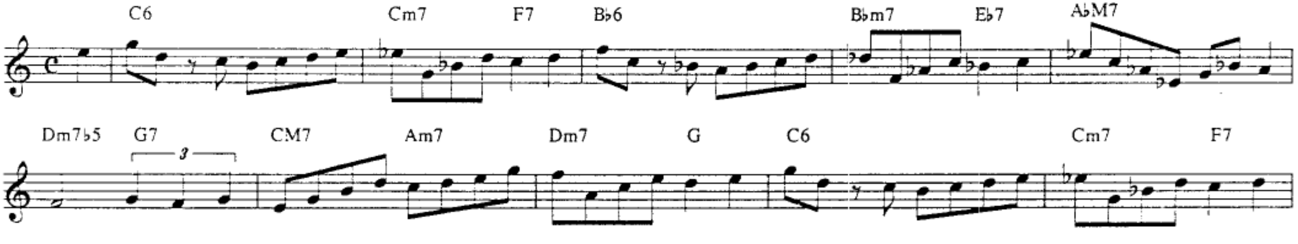
**Figure 3. First 10 bars of the Jazz standard "Afternoon in Paris".**

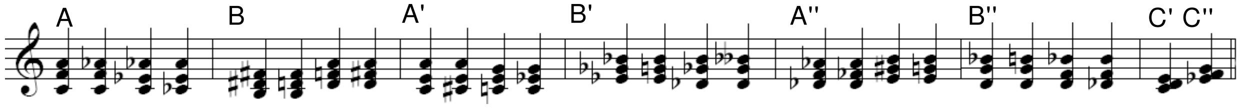

**Figure 5. Surface chords for the three different deep structure categories A, B, C in both grammars of the present study.**

prototypical sequences [B A C] or [A C] which could be recursively attached to the left of element B. The structure on the top level was chosen to be [B A] or [B A B]. In this way, hierarchically nested and recursive sequence structures such as [B A [B A C] B], or, [[[B A C] B A C] B A] could be created (as illustrated in Figure 4) that were similar in nature to the harmonic sequences depicted in Figure 3. It was decided to limit the number of these structural embeddings to a maximum of 3 layers, so that limits of participants' short-term capacities when dealing with this unfamiliar new system could be accomodated. Under this limitation sequences will not be longer than 9 elements.

In order to formalise these structures, the following context-free grammar (Chomsky, 1956) could be formulated:

(1) s → x A | x A x
(2) x → B | y B
(3) y → x A C | A C

These rules capture the structure above: the grammar contains three rules, three (lowercase) nonterminal variables s, x, y and three (uppercase) terminal symbols A, B, C. The production starts with the variable s, and using the three rules; the sequence is continuously rewritten replacing one variable within the sequence by either of the structures (separated by the '|'operator) on the right-hand side of the rule until there are no variables in left in the sequence. One production could be s → x A → y B A → x A C B A → B A C B A using the rules (1), (2), (3), (2) in this order. Starting a sequence from the variable s, the three rules produce the embedded structure as outlined. Rule (2) produces the symbol B with or without the additional attached sequence y, which represents the [B A C] or [B A] structure, attached to its left. Rule (3) defines y as the [B A C] or [B A] structure, in which by virtue of rule (2), B could have another structure attached or not.

Rules (2) and (3) are recursive because a rewrite step of the variable x using rule (2) may produce the variable y and another subsequent rewrite of the variable y may produce x again. However there is no infinite recursion since there are rewrite options for the variables without reproducing x or y. With regard to the top level, the structures may produce long-distance dependencies when the first layer [B A B] is intermitted as [B A […] B].

In order to model the key shifting and embedding illustrated in Figure 3, the surface structure for the system was designed in a way that each instance of the sequence [B A C], as in the sample sequence [B [[B A C] B A C] A B], would be transposed by a certain interval in order to create similar structural relationships. As the number of embeddings was limited to three, a realisation using an octatonic system was created: since there are only 3 different transpositions of the octatonic scale, the surface chord elements for each [B A C] sequence could be taken from the same octatonic scale so that the process of embedding would entail a change of scale and a respective transposition. Each octatonic scale contains 4 major and 4 minor chords as well as 8 diminished triads. It was chosen to use 4 of these 8 major and minor chords as possible surface exemplars of each class A and B. The remaining class C was created by employing dissonant triads in order to avoid association with tonal harmony. It was possible to find a solution in which sets of 4 chords for A and B could be selected to embody an as minimal as possible number of fifth relationships between adjacent chords in the resulting surface sequences, and thus to ameliorate associations to tonal harmony (Figure 5). Using the octatonic scale on F, the set of {F, Fm, Ab, Abm} chords were used for A, and the set of {B, Bm, D, Dm} for B for the first (top) layer. Using this assignment, a change of the octatonic scale for an embedding would result in a transposition by a major third: A' and B' (for the second embedded layer) would be {A, Am, C, Cm} and {Eb, Ebm, Gb, Gbm}, and for the third embedding {Db, Dbm, E, Em} and {G, Gm, Bb, Bbm} for A'' and B''. This way, all 24 major and minor chords were included without overlap in this new system. For the C chords that would happen in the two embedded layers only, the dissonant chords C-D-E and E-Gb-Ab were chosen for C' and C'' respectively. An embedded sequence of [B A [B' A' C'] B] could therefore have, for example, the following surface representation (Figure 6): [B Fm [Ebm C (C-D-E)] D]. In a sequence like this, the major and minor chords are used entirely out of the context of their tonal functions, so that minimal relationships were drawn to tonal harmonic structure. [1]
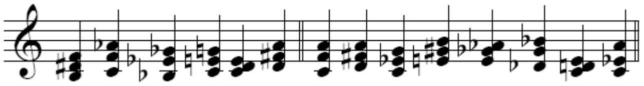
**Figure 6. Examples of the surface realization for B A [B A C] B (grammar 1, left) and B A [A [A C] B C] B (grammar 2, right)**

In order to acquire knowledge about the underlying deep structure of the relationships between the three classes of chords (A, B, C) in such sequences, the participants would have to find a representation of the octatonic relationships as well as the similar chord relationships within the embedded sequences. That is, the participants would need to acquire from the number of examples that chords from the sets of A and B and B and C frequently co-occur.

---

[1] One may notice that the scale transposition that was added when creating the surface structure from the context-free deep structure has an influence on the inferable grammar. However, when using phrase-structure rules with variables to model transposition (as discussed by Steedman, 1996; or as featured in lexical-functional grammar, Bresnan, 2001) the entire structure can be expressed in an analogous context-free way. It was decided, however, to omit a description of this all-encompassing formalism due to the space limitations and in order to avoid additional complexity.

# III. EXPERIMENT I

## A. Method

*1) Participants.* Due to the considerable complexity of the structure, it was decided to use only musicians in this study. 18 adults (9 men, 9 women, mean age 20.7 years) and native speakers of English participated in the study. The participants featured high experience and training in Western classical music (14.3 years of formal musical training on average) and average practical music making of 14.4 hours per week.

*2) Stimulus materials.* The grammar featured the three rules and the sequence construction principles as outlined above. Applying the above constraint that there were no more than three levels of embedding, the grammar produces 18 different abstract structures of the length between 2 and 9 chords. Only 10 of these 18 structures were selected for the learning phase (old- grammatical), and the 8 remaining structures (new-grammatical) as well as the 10 old-grammatical structures were both used in the testing phase.

Exploiting all possibilities to insert one of the possible terminal surface chords for each of the 3 grammatical chord types involved, the 18 grammatical structures create several thousand combinatorial possibilities. For the learning phase of the experiment, 168 surface sequences for the 10 old-grammatical structures were randomly chosen so that there were 16 structures with one layer, 80 structures with two layers and 72 structures with three layers. It was decided to present the short 1-layer sequences to the participants as well in order to familiarise them well with the top-level structure.

For the testing phase, 44 new surface exemplars were randomly chosen for the old-grammatical structures as well as 22 new surface exemplars for the remaining new-grammatical structures. In total, 66 grammatical stimuli were randomly selected for the testing phase.

Ungrammatical stimuli of different kinds were created in order to be able to explore the extent to which participants acquired specific features of the grammar. There were 22 random stimuli (sampled from the unigram chord distribution of the grammatical structures from the learning phase) in order to provide a baseline showing whether participants were able to distinguish grammatical from mere random structures. In addition, there were ungrammatical stimuli that featured systematic violations of the embedded sequences. There were 6 stimuli with 3 layers in which layer 1, 2, 3 was scrambled, 10 stimuli with 2 layers in which layer 1 and 2 was scrambled and 10 stimuli with one layer in which this layer was scrambled.

For the presentation during the test, every ungrammatical stimulus was matched with a grammatical stimulus of the same but correct abstract structure. Table 1 summarises the organisation of the stimuli.

All stimuli were computationally generated using the MIDI Toolbox for MATLAB (Eerola & Toiviainen, 2004). For the generation, the chord sequences were limited to the range of one octave in order to limit avoid unnatural leaps between chords, and to create a condition that would guarantee that adjacent chords would have a close voicing without imposing an additional set of voice-leading rules.

Then, MIDI files using a piano timbre were created using the MIDI Toolbox and subsequently rendered to audio WAV format using WINAMP. All stimuli were played over headphones.

**Table 1. Numbers of stimulus structures for Experiment 1**

| | **Grammatical** | | **Ungrammatical** | | | |
|---|---|---|---|---|---|---|
| | **Old-gr** | **New-gr** | **Random** | **Viol. L1** | **Viol. L2** | **Viol. L3** |
| 1 Layer | 12 | | 6 | 6 | | |
| 2 Layers | 15 | 15 | 10 | 10 | 10 | |
| 3 Layers | 12 | 12 | 6 | 6 | 6 | 6 |
| Total | 39 | 27 | 22 | | 44 | |

*3) Procedure.* During the learning phase, which lasted about 25 minutes, participants listened to 168 stimuli and were ask to perform a distraction task of counting the number of chords in each sequence. The subsequent testing phase, which lasted about another 30 minutes, employed Jacoby's (1991) process dissociation procedure and consisted of an inclusion and an exclusion task of 66 trials each. The inclusion task required participants to choose the familiar one of two stimuli (of identical length) using a forced-choice 2-alternative paradigm and was followed by a confidence rating on a 6-step scale of

**Table 3. Mean standardised beta coefficients for the logistic regression analysis exploring the extent to which grammaticality, bigrams and local repetition structure predict participants' responses**

| | **Experiment 1** | | | | | | **Experiment 2** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Inclusion** | | | **Exclusion** | | | **Inclusion** | | | **Exclusion** | | |
| | **M** | **t(18)** | **sig** | **M** | **t(18)** | **sig** | **M** | **t(17)** | **sig** | **M** | **t(17)** | **sig** |
| **Layer violations** | | | | | | | | | | | | |
| Intercept (grammaticality) | 0.719 | 6.316 | 0.000 | 0.307 | 2.340 | 0.031 | 0.963 | 8.101 | 0.000 | 0.682 | 4.423 | 0.000 |
| Bigrams | 0.017 | 1.950 | 0.067 | 0.027 | 2.752 | 0.013 | 0.021 | 2.387 | 0.029 | 0.002 | 0.144 | 0.888 |
| Local repetition | 0.008 | 2.086 | 0.051 | 0.005 | 1.652 | 0.116 | -0.010 | -3.171 | 0.006 | -0.007 | -2.356 | 0.031 |
| **Random structures** | | | | | | | | | | | | |
| Intercept (grammaticality) | 3.058 | 1.526 | 0.144 | 2.200 | 1.256 | 0.225 | 9.511 | 2.640 | 0.017 | 3.565 | 1.718 | 0.104 |
| Bigrams | 0.014 | 0.426 | 0.675 | -0.021 | -0.283 | 0.780 | 0.054 | 0.548 | 0.591 | 0.141 | 1.773 | 0.094 |
| Local repetition | 0.054 | 6.111 | 0.000 | 0.079 | 2.131 | 0.047 | -0.002 | -0.061 | 0.952 | 0.001 | 0.024 | 0.981 |

50%-100%. (Dienes & Longuet-Higgins, 2004). The exclusion task required people to choose the unfamiliar item to two stimuli using forced-choice 2-alternative responses and subsequent confidence ratings as before. The experiment concluded with a debriefing questionnaire applying subjective measurements of implicit knowledge (Dienes & Scott, 2005) in which participants were asked about their intuitions about an underlying rule system and potential response strategies they had used.

### B. Results

Table 2 lists the results. Planned one sample t tests against a 0.5 chance level revealed that participants performed significantly above chance in terms of distinguishing between random and grammatical structures and in distinguishing grammatical structures from systematic layer violations under both, inclusion and exclusion conditions. Participants also performed above chance for all 3 types of layer violations and for both local and long-distance dependencies. The performance for old-grammatical structures was not statistically different from the performance for new-grammatical structures, t(17)=1.01, p=.326, t(17)=1.30, p=.212 for inclusion and exclusion respectively. In general, these results suggests that participants have acquired some generalised knowledge that enables them to distinguish grammatical from ungrammatical structures independently of the deep structures used in the training examples.

## IV. EXPERIMENT II

The purpose of Experiment 1 was to explore whether participants were able to acquire some features of embedded hierarchical structure from mere exposure. A subsequent experiment was performed to investigate whether there was an effect of grammatical complexity if the grammar used was altered to be more complex. Whereas the grammar used in Experiment I generated centre-embedded structures only with respect to the first layer being interpolated, a more complex structure would also allow the creation of centre-embeddings and long-distance dependencies on all levels. In order to create such a change in complexity a small change in rule (3) of the grammar to rule (3') is sufficient:

(3') y → A x C | A C
from (3) y → x A C | A C

This small manipulation changes the order of the sequence [B A C] to [A B C]. However, since another example of the embedded sequence may be attached to the left of B, now the embedding constitutes a centre-embedded structure and the structure of [A B C] will become [A [A B C] B C] when interpolated. Therefore, the formerly paratactic sequences of [B A [[B C] B A C] B] or [[[A C] B A C] B A] become [B A [A [B C] B C] B] or [[A [A C] B C] B A].

We hypothesised that an increase of complexity would have a negative effect on participants' performance, due simply to the fact that more complex structures might be more difficult to learn or to recognise.

Whereas there is no genuine difference between a more complex phrase-structure grammar and a more simple one – both systems are an equally plausible structure within the formalism – one might expect both structures to be comparably frequent in a context-free system. However, a cursory glance through a book of Jazz standards or classical harmony finds some examples for the first structure but considerably fewer examples for the second, more complex, structure. So far, there is no empirical evaluation of this difference, even though work is in preparation (de Haas & Rohrmeier, in preparation). Based on Hawkins's (2004) theory that aspects of efficiency and complexity shape the way how structures are produced, reproduced and transmitted throughout the course of time, human cognitive constraints favour the use of structures which can be processed efficiently and have a lower degree of complexity. Therefore, syntactic structures that feature more efficient communication and lower degrees of complexity would tend to dominate more throughout time than more complex structures. Hawkins outlines that comparable linguistic corpus studies have found syntactical structures such as in grammar 1 to be more frequent across the languages of the world than the more complex grammar 2 (Dryer, 1980). The aim of this study was not to explore Hawkins's hypothesis in musical terms, but to link with it by exploring to which extent there is an effect of grammatical complexity in learning and recognising structures of different complexity.

### A. Method

*1) Participants.* 19 adults (11 man, 8 women, mean age 20.7 years) and native speakers of English participated in the study. As in Experiment 1, the participants had high experience and training in Western classical music (13.8 years of formal musical training in average) and average practical music making of 12.7 hours per week.

*3) Stimuli.* The new grammar produced an identical number of 18 abstract structures. Hence, the stimuli were organised the same way as in Experiment 1 and used an identical rationale for the divisions into old- and new-grammatical structures, and numbers of grammatical and ungrammatical stimuli was applied. The same constraint of up to 3 layers of embedding was used, with a maximal sequence length of 9 chords was applied. The same octatonic system and stimulus creation method as in Experiment 1was also used to create new sets of grammatical (see Figure 6 for an example chord sequence) and layer-violating structures. The same set of random stimuli was used.

*4) Procedure.* The procedure was identical to the procedure in Experiment I.

### B. Results

Table 2 lists the results. Planned one sample t tests against a 0.5 chance level revealed that participants performed significantly above chance in terms of distinguishing between random and grammatical structures and in distinguishing grammatical structures from systematic layer violations under both, inclusion and exclusion conditions. In addition, participants performed above chance for all 3 types of layer violations (except layer 3 violations under exclusion) and for both, local and long-distance dependencies. The performance for old-grammatical structures was higher than for new-grammatical structures under inclusion, t(18)=2.44, p=.025, but not under exclusion t(18)=1.57, p=.134. Altogether these findings suggest that participants had acquired some knowledge that enabled them to distinguish grammatical from ungrammatical structures.

## V. COMPARING BOTH EXPERIMENTS

The results show significant above-chance performance in both experiments. Post hoc ANOVAs explored the extent to which there was an effect of grammatical complexity between both experiments.

An ANOVA with task (inclusion vs. exclusion) and type of ungrammatical type (random ungrammatical vs. layer-violations) as within-subject variables and group as between-subject variable indicated a significant effect of group, F(1,35)=6.26, p=.017, a highly significant within-subject effect of ungrammatical type, F(1,35)=186.14, p<.0005 and a highly significant effect of task, F(1,35)=21.23, p<.0005. This indicates that participants in both groups were better at distinguishing grammatical from random structures than from systematic layer violations.

An ANOVA with task (inclusion vs. exclusion) and type of layer violation (layer 1 vs. layer 2 vs. layer 3) as within-subject variables and group as between-subject variable indicated no significant effect of group, a significant within-subject effect of task, F(1,35)=23.60, p<.0005, a significant interaction between layer-violation and task, F(1.72,35)=6.87, p=.003, and a significant within-subject effect of layer-violation, F(1.72,35)=4.09, p=.027 using the Greenhouse-Geisser correction in both cases. Simple contrasts comparing the performance for structures violating layer 2 or 3 with structures violating layer 1 indicated a significant difference between the performance for layer 1 and layer 2 violations across both groups (F(1,35)=10.60, p=.03) and an interaction with group with respect to the difference between layer 1 and layer 3 violations (F(1,35)=11.46, p=.02). This suggests that participants' performance differed for the violations of different layers and varied with grammatical complexity. In both experiments layer 2 violations were the most difficult to detect. In the case of the more complex grammar, layer 3 violations (which featured the deepest embedding) were more difficult to detect than layer 1 violations, whereas for the simpler grammar layer 3 violations were in fact better detected than layer 1 violations.

An ANOVA with task (inclusion vs. exclusion) and dependency type (long-distance dependencies vs. local dependencies) as within-subject variables and group as between-subject variable indicated a significant effect of group, F(1,35)=6.83, p=.013, a significant interaction of dependency type and group, F(1,35)=5.71, p=.022, and a significant within-participants effect of task, F(1,35)=23.67, p<.0005. This indicates that the performance difference for local vs. long-dependencies was different between both experiments.

### C. Form of representation

One central question with respect to the implicit learning literature concerns what type of knowledge and features participants have acquired which enables them to recognise the structures above chance. It would be naive to argue that participants had acquired the abstract grammatical structure behind the stimuli (Reber, 1993). For example, it is known that participants acquire and use knowledge of small fragments of the surface structure (such as bigrams or trigrams), and that participants are sensitive to features of local element repetition in single stimuli (Meulemans & Van der Linden, 1994). Therefore, it is unclear whether they acquired features of the grammatical structure, deep structure, or some surface fragments or regularities. We therefore used a logistic regression method , common in implicit learning (Dienes & Longuet-Higgins, 2004; Kuhn & Dienes, 2005, 2006) to determine the extent to which grammatical structure, surface bigrams, and local repetition structure predicted participants' responses.

Following this method, for each participant a logistic regression was carried out that measured the degree to which grammaticality, bigram structure, or local repetition structure predicted the participant's responses for the set of stimuli. Bigram matches were calculated according to the method by Meulemans & Van der Linden (1994) by counting for each stimulus in the testing set how often each of its bigram fragments occurred in the training set. Hence, large numbers would indicate that a simulus contains many bigram fragments that occurred during learning, and small numbers would indicate the opposite. Surface bigram and trigram matches were strongly correlated (r=0.84). Therefore only bigrams were used for the regression in order to maintain solution stability. Local repetition structure (Meulemans & Van der Linden, 1994) was coded by calculating for each testing stimulus how often a stimulus with the same repetition pattern between adjacent surface chords (for example, a local repetition structure of 000110, indicates that in a sequence of 7 chords the first 4 are different, chords 4,5,6 are identical and chords 6 & 7 are different) occurred in the learning phase.

The logistic regression was performed separately for inclusion and exclusion tasks as well as for layer violations and random structures. Table 3 displays the results. Regarding the analysis of systematic layer violations, it was found that for both inclusion and exclusions tasks in Experiment 1 grammatical structure was a significant and the strongest predictive factor for participant responses: m=0.72, t(18)=6.32, p<0.0005 and m=0.31, t(18)=2.34, p=.031 respectively. In the exclusion case, bigrams turned out as a significant predictor even though the mean value was small compared to grammatical structure: m=0.03, t(18)=2.75, p=.013. Similarly, grammatical structure was the strongest predictor in Experiment 2: m=0.97, t(17)=8.10, p<.0005 and m=0.69, t(17)=4.423, p<.0005 for inclusion and exclusion respectively. Under inclusion, surface bigrams had a small, yet significant effect: m=.02, t(17)=2.39, p=.029, and local repetition structure had a small, yet significant, negative effect: m=-0.01, t(17)=-3.17, p=.006 and m=-.007, t(17)=-2.36, p=.03 respectively. Overall, this suggests that grammatical structure was the best predictor of participants' responses across both experiments and both tasks. In two cases, surface bigrams were found to be a minor predictor of participant responses and local repetition structure had a small negative (i.e. detrimental) effect.

In the cases of random structures, local surface structure was a significant predictor across participants (m=0.05, t(18)=6.11, p<.0005 and m=0.08, t(18)=2.13, p=.05 for inclusion and exclusion respectively), indicating that participants used, in part, repetition features to determine their responses. Despite very high mean values (m=3.06, m=2.20 respectively) grammatical structure had no significant effect on responses across participants in Experiment 1 under inclusion and exclusion, indicating that for some participants grammatical structure was a strong predictor whereas it was not for others. In Experiment 2, however, local repetition structure yielded no significant patterns at all, and grammatical structure was found to be a significant predictor (m=9.51, t(17)=2.64, p=.017) under inclusion whereas none of the three predictors were significant across participants under exclusion suggesting that there were no consistent common response patterns in that case, even though the mean value for grammatical structure is, again, very high (m=3.57) compared to the other predictors.

Overall, grammatical structure had a strong and significant impact in layer-violation cases for both tasks in both experiments. This suggests that participants acquired some representations beyond mere surface fragments that they were able to match to features of the grammatical structure. However, it is impossible to reveal whether participants learned the actual grammar or another representation (Reber, 1993; Dienes & Longuet-Higgins, 2004).

Following on from that, a final analysis was carried out in order to explore whether there was an effect of grammatical complexity once surface features such as bigrams and local repetition structure were controlled for. An ANOVA with instruction (inclusion vs. exclusion) as within-subject variable and complexity as between-subject variable based on mean regression coefficients for grammatical structure indicated a significant effect of complexity, F(1,35)=4.73, p=.36, and a significant effect of instruction, F(1,35)=8.837, p=.005. Therefore, grammatical complexity may be assumed to have influenced the learning of grammatical structure.

### D. Awareness

In order to analyse whether confidence ratings suggested implicit or explicit judgment knowledge, the analysis method by Dienes & Longuet-Higgins (2004) was adopted. Linear regression analyses for responses against confidence level were calculated for each participant separately and intercepts as well as slope values were collected for each participant. One-sample t tests show that the set of intercepts under inclusion and exclusion are significantly below chance, m=0.35, t(18)=7.57, p<.0005, m=0.40, t(18)=7.60,p<.0005 for Experiment 2, m=0.39, t(17)=7.09, p<.0005, m=0.37, t(17)=7.78, p<.0005, for Experiment 1, inclusion and exclusion respectively. The slopes are significantly above chance for both experiments under inclusion and exclusion: m=0.11, t(18)=10.12, p<.0005, m=0.08, t(18)=4.53, p<.0005, for Experiment 2, m=0.11, t(17)=8.81, p<.0005, m=0.10, t(17)=8.33, p<.0005, for Experiment 1, inclusion and exclusion respectively).

These findings suggest that participants were in general significantly worse than chance when they indicated they were literally guessing, and that performance improved significantly with confidence. These findings suggest that participants acquired in part explicit judgment knowledge (Dienes & Scott, 2005), i.e. they knew to some extent when they were right.

However, the fact that participants choose grammatical structures as familiar in both inclusion and exclusion tasks above chance, in particular, under the exclusion task, which required them to select the less familiar stimulus, indicates that they also possess implicit structure knowledge (Dienes & Scott, 2005) to a high degree as implicit familiarity knowledge was overriding the opposing exclusion task instructions (Jacoby, 1991). We therefore conclude that participants acquired both implicit structure knowledge as well as explicit judgment knowledge.

The debriefing questionnaires suggest that participants were unable to articulate any specific regularities or rules of the sequences even though they would possess the necessary vocabulary being highly trained music students. One participant recognised that the sequences were organised by an octatonic system.

## VI. GENERAL DISCUSSION AND CONCLUSION

The aim of this research was to determine the learnability of harmonic structure with respect to different complexity levels, hierarchical organisation and long-distance dependencies. The results imply that participants were indeed able to learn to hear familiar chord units within a new musical system and acquire new and unfamiliar harmonic structures with complex features. Participants were able to learn to distinguish grammatical structures from ungrammatical structures for both the simpler and the more complex grammars. The fact that participants acquired knowledge about the structures, even though the learning phase used a distraction task and no explicit instructions to memorise or to learn, constitutes evidence for *incidental* learning.

The fact that the structures in the learning and testing phase were similar with regard to the deep structure, yet entirely different on a surface level, shows that participants picked up some form of knowledge which enabled them to recognise some of the grammatical sequences on the underlying deep structure level. This is supported by the fact that the logistic regression analysis revealed that surface bigram fragments or repetition structures had little power to .predict participants' responses. The additional finding that participants performed highly above-chance for both, old- and new-grammatical deep structures provides evidence for the acquisition of a more generalised form of knowledge beyond mere (deep structure) memorisation of the sequences from the learning phase and across the old-/new-grammatical distinction. When the grammar was complex, this generalisation was slightly impaired in that participants were not as good at recognising new-grammatical structures, but were still significantly above chance.

The subsequent logistic regression supports the argument that participants' responses could not easily be explained through surface fragments or repetition structures, and that grammaticality was the strongest predictor. Having controlled for the other predictors, there was a main effect of complexity which indicates that grammatical complexity had some impact on the learning and recognition of the deep structure. It remains a matter of future work to explore whether the findings will extend to nonmusicians or to subjects from different cultural backgrounds.

The analysis of the confidence ratings revealed that participants knew when they were giving correct responses, which suggests that they possessed explicit judgment knowledge. However, participants' high tendency to pick familiar grammatical items instead of ungrammatical ones in the indirect exclusion task indicates that they possessed additional implicit structure knowledge to a high extent in as much as its immediacy was overriding the explicit access (Jacoby, 1991).

Nonetheless, it is impossible simply to conclude that participants acquired or represented the grammar which was used to create the materials (Reber, 1993), as there are many rules or possible mental representations that could result in similar response patterns. However, the findings suggest that participants acquired some form of representation which enabled them to distinguish grammatical from ungrammatical structures, i.e. to deal with structures that contained hierarchical embedding and long-distance dependencies.

In this respect, the findings link to the overarching cognitive science debate in as much as they provide evidence that, in the case of this setup and grammar, participants could acquire knowledge to deal with hierarchical structure and long-distance dependencies in the domain of music. This conforms to experimental findings that musicians and nonmusicians are able to perceive long-distance dependencies and embedded structures in tonal harmony (Woolhouse et al, submitted). The findings of the present study link with Hawkins's (2004) complexity and efficiency theory, and Dryer's (1980) empirical findings about grammatical complexity: Hawkins argues that languages converge towards simpler grammars which serve for greater efficiency in terms of perception, production and communication. The result that the performance was reduced in the cases of the more complex grammar indicates that grammatical complexity affects learnability and recognition in music. This constitutes one piece of evidence that a parallel process, along the lines of Hawkins's theory, may hold for music. It remains an open question from this perspective whether empirical corpus studies in music may find similar patterns of complexity preferences in music as those found in language.

## ACKNOWLEDGMENTS

We would like to thank Guy Hayward and Matthew Woolhouse. This research was supported in part by Microsoft Research through the European PhD Scholarship Programme and the Arts & Humanities Research Council.

## REFERENCES

Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4):899–912.

Baroni, M., Maguire, S., & Drabkin, W. (1983). The Concept of Musical Grammar. *Music Analysis*, 2:175-208.

Baroni, M. Brunetti, R., Callegari, L., & Jacobini, C. (1982). A grammar for melody. Relationships between melody and harmony. In Baroni, M & Callegari, L. (Eds.), *Musical grammars and computer analysis.* Leo S. Olschki Editore: Firenze, 201-218.

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

Bresnan, J. (2001). Lexical Functional Syntax. Basil Blackwell: Oxford.

Burbat, W. (1988). *Die Harmonik des Jazz.* Bärenreiter: Kassel.

Chemiller, M. (2004). Toward a formal study of jazz chord sequences generated by Steedman's grammar. *Soft Computing,* 8:617-622.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory,* 2:113–124.

Chomsky, N. (1995). *The Minimalist Program.* MIT Press: Cambridge, MA.

Chomsky, N., & Schützenberger, M. P. (1963), The algebraic theory of context free languages. In P. Braffort and D. Hirschberg (eds.), *Computer Programming and Formal Languages*. Amsterdam, pp. 118-161.

Cleeremans, A.; Destrebecqz, A. & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, 2: 406-416.

Cook, N, (1987). The Perception of large-scale tonal closure. *Music Perception, 5*:197-206.

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: 1119-1130.

De Haas, B. & Rohrmeier, M. (submitted). Matching Methods Based on a Generative Grammar of Tonal Harmony.

Deliège, I., Mélen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real-time listening to a piece of music. *Music Perception*, 14:117-159.

Dienes, Z. & Longuet-Higgins, C. (2004). Can musical transformations be implicitly learnt? *Cognitive Science*, 28:531–558.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, 69: 338-351.

Dryer, M. (1980). The positional tendencies of sentential noun phrases in universal grammar. The Canadian Journal of Linguistics= La Revue Canadienne, 25: 123–195.

Eerola, T. & Toiviainen, P. (2004). *MIDI toolbox: MATLAB tools for music research*. Jyväskylä: University of Jyväskylä

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science, 303*: 377.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*:179–210.

Friederici, A. D., Bahlmann, J., Helm, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 2458–2463.

Gauldin, R. (1997). *Harmonic Practice in Tonal Music*. Norton, NewYork.

Giblin, I. (2008). Is Music Recursive? Poster presentation at Music, Language & the Mind conference, Tufts University, Cambridge Mass.

Gjerdingen, R. (1999). An experimental music theory?, In N. Cook & M. Everist (Eds.), *Rethinking Music*. Oxford: Oxford University Press: Oxford. 161-170.

Hawkins, J. (2004). *Efficiency and Complexity in Grammars*. Oxford University Press: Oxford.

Hochmann, J. R., Azadpour, M., & Mehler, J. (2008). Do humans really learn AnBn Artificial Grammars From Exemplars? *Cognitive Science,* 32: 1021-1036.

Jackendoff & Lerdahl (2003). The Capacity for Music: What's Special about it? *Cognition* 100: 33-72.

Jackendoff, R. (2009). Parallels and Nonparallels between Language and Music. *Music Perception, 26*: 195-204.

Jackendoff, R.S. (2007). *Language, consciousness, culture. Essays on mental structure.* Cambridge, MA: MIT Press.

Jacoby, L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30: 513–541.

Kostka, S. & Payne, D. (1984). Tonal harmony with an introduction to 20th-century music. McGraw-Hill.

Kuhn, G., & Dienes, Z. (2005). Implicit learning of non-local musical rules. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31(6):1417–1432.

Kuhn, G., & Dienes, Z. (2006). Differences in the types of musical regularities learnt in incidental and intentional learning conditions. *Quarterly Journal of Experimental Psychology*, 59: 1725-1744.

Lerdahl, F. & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA.

Levine, M. (1996). *The Jazz Theory Book*. Sher Music.

Loui, P., & Wessel, D. (2008). Learning and linking an artificial musical system: Effects of set size and repeated exposure. *Musicae Scientiae*, 12(2):207-230.

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23:1007–1028.

Parncutt, R. (1989). Harmony: A psychoacoustical approach. Springer.

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

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

Perruchet, P. & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119:264–275.

Perruchet, P., & Rey, A. (2005). Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin & Review, 12,* 307–313.

Pesetsky, D. (2007). Music Syntax is Language Syntax. Talk and Handout presented at Language and Music as Cognitive Systems Conference, Canbridge.

Pinker, S. & Jackendoff, R. (2005) The faculty of language: what's special about it? Cognition 95: 201–-236.

Piston, W. (1948). *Harmony*. W.W.Norton & Company, New York.

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin*, 133(2):227–244.

Reber, A. S. (1967). Implicit learning of artifical grammar. *Journal of Verbal Learning & Verbal Behaviour*, 6:855–863.

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

Rohrmeier, M., Fu. Q., & Dienes, Z. (in prep.) Implicit learning of recursive, hierarchical grammatical structures.

Rohrmeier, M., Rebuschat, P. & Cross, I. (submitted). Incidental learning of melodic structure.

Rohrmeier, M. (2007). A generative grammar approach to diatonic harmonic structure. In Spyridis, Georgaki, Kouroupetroglou, Anagnostopoulou (Eds.), *Proceedings of the 4th Sound and Music Computing Conference*, 97-100.

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

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

Schellenberg, E.; Bigand, E.; Poulin, B.; Garnier, C., & Stevens, C. (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science*, 8: 551-566.

Schenker, H. (1935). Der Freie Satz. Neue musikalische Theorien und Phantasien. Liege: Margada.

Shanks, D. R. (2005). Implicit learning, In Lamberts, K. & Goldstone, R. (Eds.), *Handbook of Cognition,* Sage, 202-220.

Steedman, M. (2002). Plans, affordances, and combinatory grammar. *Linguistics and Philosophy*, 25: 723–753.

Steedman, M. (1996). The Blues and the Abstract Truth: Music and Mental Models, In A. Garnham & J. Oakhill (Eds.), *Mental Models In Cognitive Science*. Erlbaum, Mahwah, NJ, 305-318.

Steedman, M. (1984). A Generative Grammar for Jazz Chord Sequences. *Music Perception* 2: 52-77.

Tekman, H. G. & Bharucha, J. (1992). Time course of chord priming. *Perception & Psychophysics,* 51: 33-39.

Thompson, W.F. (2009). *Music Thought & Feeling: Understanding the Psychology of Music*, Oxford University Press: New York

Tillmann, B. & McAdams, S. (2004). Implicit learning of musical timbre sequences: Statistical regularities confronted with acoustical (dis)similarities. *Journal of Experimental Psychology,* 30: 1131-1142.

Tillmann, B. (2005). Implicit investigations of tonal knowledge in nonmusician listeners (pp. 100-110). In : The Neurosciences and Music II: From Perception to Performance, Volume 1060 of the Annals of the New York Academy of Sciences, edited by Giuliano Avanzini, Luisa Lopez, Stefan Koelsch, & Maria Majno.

Tojo, O. Y. & Nishida, M. (2006). Analysis of chord progression by HPSG. In *Proceedings of the 24th IASTED international conference on Artificial intelligence and applications*, 305-310.

Tomalin, M. (2007). Reconsidering recursion in syntactic theory. *Lingua*, 117: 1784-1800.

Woolhouse, M., Cross, I. & Horton, T. (submitted). The perception of non-adjacent harmonic relations.

**Table 2. Participants' performance and one-sample t tests against chance level for both experiments. $^{*}$: p<.05 $^{**}$: p<.005 $^{***}$: p<.0005**

| | **Experiment 1** | | | | | | **Experiment 2** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Inclusion** | | | **Exclusion** | | | **Inclusion** | | | **Exclusion** | | |
| | **Mean** | **SD** | **t(18)** | **Mean** | **SD** | **t(18)** | **Mean** | **SD** | **t(17)** | **Mean** | **SD** | **t(17)** |
| All layer violations | $0.679^{***}$ | 0.083 | 9.389 | $0.606^{**}$ | 0.116 | 3.995 | $0.740^{***}$ | 0.089 | 11.398 | $0.654^{***}$ | 0.099 | 6.596 |
| Random structures | $0.830^{***}$ | 0.084 | 17.102 | $0.751^{***}$ | 0.166 | 6.614 | $0.891^{***}$ | 0.083 | 20.079 | $0.841^{***}$ | 0.094 | 15.408 |
| Violation Layer 1 | $0.684^{***}$ | 0.109 | 7.387 | $0.598^{**}$ | 0.102 | 4.194 | $0.793^{***}$ | 0.132 | 9.380 | $0.694^{***}$ | 0.145 | 5.679 |
| Violation Layer 2 | $0.645^{***}$ | 0.133 | 4.726 | $0.582^{*}$ | 0.167 | 2.150 | $0.667^{***}$ | 0.115 | 6.125 | $0.632^{***}$ | 0.100 | 5.581 |
| Violation Layer 3 | $0.754^{***}$ | 0.170 | 6.521 | $0.702^{***}$ | 0.205 | 4.296 | $0.741^{***}$ | 0.164 | 6.231 | 0.565 | 0.199 | 1.381 |
| Old grammatical | $0.757^{***}$ | 0.073 | 15.412 | $0.671^{***}$ | 0.123 | 6.033 | $0.801^{***}$ | 0.097 | 13.098 | $0.729^{***}$ | 0.099 | 9.809 |
| New grammatical | $0.690^{***}$ | 0.115 | 7.227 | $0.632^{**}$ | 0.143 | 4.021 | $0.776^{***}$ | 0.089 | 13.092 | $0.698^{***}$ | 0.097 | 8.644 |
| Local dependencies | $0.704^{***}$ | 0.120 | 7.396 | $0.618^{**}$ | 0.152 | 3.403 | $0.793^{***}$ | 0.110 | 11.295 | $0.725^{***}$ | 0.105 | 9.120 |
| Long-dist. dependencies | $0.744^{***}$ | 0.079 | 13.499 | $0.675^{***}$ | 0.129 | 5.939 | $0.786^{***}$ | 0.070 | 17.333 | $0.703^{***}$ | 0.086 | 9.978 |