TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 4 (2012) 525–553

ISSN: 1756-8757 print / 1756-8765 online
DOI: 10.1111/j.1756-8765.2012.01223.x

Implicit Learning and Acquisition of Music

Martin Rohrmeier,[a] Patrick Rebuschat[b]

[a]*Centre for Music & Science, Faculty of Music, University of Cambridge, and Cluster Languages of Emotion, Freie Universität Berlin*
[b]*School of Linguistics and English Language, Bangor University*

Received 11 January 2011; received in revised form 21 February 2012; accepted 26 April 2012

Abstract

Implicit learning is a core process for the acquisition of a complex, rule-based environment from mere interaction, such as motor action, skill acquisition, or language. A body of evidence suggests that implicit knowledge governs music acquisition and perception in nonmusicians and musicians, and that both expert and nonexpert participants acquire complex melodic, harmonic, and other features from mere exposure. While current findings and computational modeling largely support the learning of chunks, some results indicate learning of more complex structures. Despite the body of evidence, more research is required to support the cross-cultural validity of implicit learning and to show that core and more complex music theoretical features are acquired implicitly.

Keywords: Music; Implicit learning; Statistical learning; Musical acquisition; Incidental learning; Computational modelling; Cognitive modelling; Music cognition

1. Introduction

Music, like language, is a historically evolved, complex, and highly structured form of human interaction and communication, which involves a range of cognitive processes in perception, parsing, and production. Basic tasks of music cognition like prediction, melody recall and comparison, style recognition, or even the ability to recognize (stylistic) mistakes in a musical excerpt rely on knowledge of the syntax governing any musical style, be it Western tonal music from Monteverdi to the Beatles, Middle Eastern maqam, traditional North Indian music, Western, Non-Western or early modal music, or Jazz improvisation. Musical errors such as ''sour'' chords or notes strike us instantly, even if we cannot explain why a given sequence sounds wrong. We are able to recognize pieces of music or the style of a piece in

Correspondence should be sent to Martin Rohrmeier, Centre for Music & Science, Faculty of Music, University of Cambridge, UK, and Cluster Languages of Emotion, Freie Universität Berlin, Germany. E-mail: mrohrmeier@cantab.net

fractions of a second (Dalla Bella, Peretz, & Aronoff, 2003), and we possess intuitions about music irrespective of whether we have received any formal musical training. Most members of a society are competent music listeners (Bigand & Poulin-Charronnat, 2006; Honing, 2009) even when they have had little or no formal musical training (Bigand & Poulin-Charronnat, 2006; Koelsch, Gunter, & Friederici, 2000). Likewise, music theorists may not be able to explain some of their intuitions explicitly. Musical knowledge, like native-language knowledge, is implicit; it is mentally represented without awareness of the entirety of the complex grammar of rules, having been acquired through attending and interacting with a large number of samples. The questions of how we manage to acquire implicit musical knowledge merely from exposure and interaction with, mostly positive, examples and of how we represent it guide the research in this domain and form a pillar of music cognition.

The process of implicit learning, generally defined as the ability to acquire unconscious knowledge without intending to, is an elementary and ubiquitous process of human cognition (see Cleeremans, Destrebecqz, & Boyer, 1998; Dienes, 2011; Perruchet, 2008; Reber, 1989, 1993; Shanks, 2005; Williams, 2009, for overviews). Essential skills, including language comprehension and production, social interaction, and intuitive decision-making, are largely dependent on implicit knowledge (Berry & Dienes, 1993; Reber, 1993)—and the same holds for music. Implicit learning underpins the long-term acquisition of stylistic knowledge, musical enculturation, and acquisition, in both children and adults, nonmusicians and musicians (proposed by Francès, 1958/1988; Krumhansl, 1990; see also Deliège & Sloboda, 1995; Huron, 2006; Stalinski & Schellenberg, in press; Trehub, 2006; Trehub & Hannon, 2006). It underlies the rapid, exposure-driven acquisition of structure observed in both statistical and implicit learning experiments (for the distinction between both, see below) and predicts the online (real-time) learning of musical features during performances.

Unlike in language (Chomsky, 1957), however, there are few reasons to assume innate grammatical structures in the case of music, given that musical structure is not overwhelmingly complex (cf. Jackendoff & Lerdahl, 2006). In fact, there is a growing body of evidence that musical competence is largely shaped by a tacit knowledge base that is acquired through exposure (Bigand & Poulin-Charronnat, 2006; Ettlinger, Margulis, & Wong, 2011; Deliège, Mèlen, Stammers, & Cross, 1996; Deliège, 2001; Huron, 2012; however, see Marcus, 2012; Hannon & Trainor, 2007; Trainor & Corrigall, 2010; for a discussion on innate and acquired musical processes). It governs participants' expectancies in priming experiments (see Huron, 2006; Tillmann, 2005) and tonal knowledge found in neuroscientific studies (Koelsch, 2011; Koelsch et al., 2000). The familiarity of implicitly acquired structures further determines the liking of such structures (Zajonc, 2001). The role of implicit knowledge in music cognition, however, exceeds the mere formation of expectancies. It underlies the perception and cognition of style-specific structures and schemata (or their limitation, Hannon & Trehub, 2005a,b), parsing processes including revision and ambiguity resolution (Jackendoff, 1991; Temperley, 2001), recognition, segmentation, disambiguation, as well as musical performance, production, and interaction between musicians. Despite its fundamental role, the absence of ''implicit learning'' and ''implicit knowledge'' in important state-of-the-art volumes such as the *Oxford Handbook of Music Psychology* (Hallam, Cross, & Thaut, 2008) suggests that the relevance of implicit learning and knowledge has not been fully recognized in the area. Although there are a number of studies of statistical and implicit learning of music, many musical features remained unexplored. This article reviews the state of implicit learning research in music cognition and corresponding evidence from computational modeling.

2. What is implicit learning?

The term *implicit learning* was first employed by Reber (1967) to describe a process during which subjects acquire knowledge about a complex, rule-governed stimulus domain without intending to and without becoming aware of the knowledge they have acquired. These two dimensions, intentionality and awareness, are central to the notions of implicit and explicit learning in the psychological literature. In the research tradition started by Reber (1967), the use of the term *implicit* is generally restricted to those situations where subjects have acquired unconscious (implicit) knowledge under incidental learning conditions. If incidental exposure in an experiment results in conscious (explicit) knowledge, for example, when subjects were able to figure out the rule system despite not having been told about its existence, the learning process is usually only characterized as being incidental and not as implicit. The same applies for those experiments that do not include a measure of awareness. The term *explicit learning* is usually applied to learning scenarios in which subjects are instructed to actively look for patterns, that is, learning is intentional, a process which tends to result in conscious knowledge.

Methodologically, we can broadly distinguish two types of studies, each following a different, although related, tradition. Studies following in the implicit learning tradition (Reber, 1967) expose subjects to musical sequences generated by a finite-state grammar. Subjects are first exposed to the stimuli under incidental learning conditions. They are then tested to determine whether learning took place and whether exposure resulted in conscious or unconscious knowledge. Studies in the statistical learning tradition (Rebuschat & Williams, 2012; Saffran, Aslin, & Newport, 1996) also employ artificial sequences to investigate the acquisition of structure. However, in contrast to research in the former tradition, statistical-learning experiments also involve the careful manipulation of statistical information in the input, for example, the manipulation of co-occurrence frequency. Statistical-learning studies mostly do not assess the conscious or unconscious status of the acquired knowledge. The terms *implicit learning* and *statistical learning* are frequently used interchangeably, and it has been argued that they refer to the same phenomenon (Perruchet & Pacton, 2006). In this review, we use the term *implicit learning* as the umbrella term embracing both directions of research and discuss the findings of the respective tradition when required.

The past decades have resulted in a relative consensus on several characteristics of implicit learning and knowledge. For example, Reber's (1967) basic finding that subjects can exploit the structure inherent to the stimulus environment incidentally has been frequently replicated and appears very robust. It is widely accepted that implicit learning causes a sense of intuition, that is, ''people do not feel that they actively work out the answer'' but rather make ''particular responses because they 'feel' right'' (Berry & Dienes, 1993, p. 14)—a finding analogous to the Mere Exposure Effect (Zajonc, 1968, 2001). Implicit knowledge seems to be more robust in the face of neurological disorder (e.g., Knowlton, Ramus, & Squire, 1992) and in the face of time, that is, it appears longer lasting than explicit knowledge (Allen & Reber, 1980). In contrast to explicit learning, successful implicit learning is less affected by individual differences (Reber, Walkenfeld, & Hernstadt, 1991), and it also seems to place fewer demands on attentional resources (e.g., Dienes & Scott, 2005). Gómez (2007) provides an extensive review of what can and cannot be acquired by means of statistical learning.

2.1. How is implicit knowledge represented?

There is ample evidence that subjects can rapidly derive information from the environment, yet how the acquired knowledge is mentally represented remains subject to debate. The original assumption (e.g., Reber, 1967, 1969) was that exposure to the artificial system results in the acquisition of abstract, rule-based knowledge, but this view has been frequently challenged and a variety of alternative explanations have been proposed over the years (e.g., Dienes, 2009; Dulany, Carlson, & Dewey, 1984; Perruchet & Pacteau, 1990; Reber, 1989, 1993; Servan-Schreiber & Anderson, 1990; see Pothos, 2007; for a review). For example, fragment-based accounts (Dulany et al., 1984; Perruchet & Pacteau, 1990; Redington & Chater, 1996; Servan-Schreiber & Anderson, 1990) suggest that subjects (merely) accumulate knowledge of fragments or ''chunks,'' such as bigrams or trigrams, weighted by their frequencies of occurrence over the training set. This chunking hypothesis further suggests that, in the testing phase, subjects classify novel test strings as grammatical to the extent to which a test string contains fragments that were present in the training strings. That is, familiarity rather than grammaticality might control performance (Johnstone & Shanks, 2001). The current cognitive debate focuses on the question—which structures beyond chunks can or cannot be learned implicitly. Although it is unsurprising that people learn context-free grammars explicitly—after all, this is what all programmers do when they learn a programming language like C or Lisp—our ability to acquire this type of structure implicitly from mere exposure is far less clear, and the same question transfers to the case of music cognition. Recent music experiments on the acquisition of cross-serial dependencies (see below) and of context-free or context-sensitive structures (cf. M. Rohrmeier, Fu, & Dienes, in press; Jiang et al., 2012) add to this ongoing debate, as the learning of such structures cannot easily be accounted for by chunking theories. The substantial complexity of musical structure provides a rich set of candidate structures for this general debate.

2.2. How ''implicit'' is implicit knowledge?

It is widely accepted that the knowledge acquired during standard implicit learning experiments is the result of an incidental learning process. However, the question of whether this knowledge is actually ''implicit'' is highly controversial (e.g., Berry, 1997; Stadler & Roediger, 1998). The case for implicit cognition depends on both our definition of awareness and on the validity of tests used to distinguish conscious (explicit) and unconscious (implicit) knowledge. Below, we will review three basic ways of distinguishing implicit and explicit knowledge (for more comprehensive reviews, see, e.g., Dienes & Berry, 1997; Gaillard, Vandenberghe, Destrebecqz, & Cleeremans, 2006; Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008). For alternative definitions of implicit cognition, see Jacoby (1991), Merikle and Joordens (1997), and Greenwald (1992).

One of the most common procedures is to prompt subjects to verbalize any patterns or rules they might have noticed while performing the experimental tasks (e.g., Dienes, Broadbent, & Berry, 1991; Lewicki, Hill, & Bizot, 1988; Reber, 1967; see Ericsson & Simon, 1980; Nisbett & Wilson, 1977; Payne, 1994; for reviews). This is generally done during the debriefing session, that is, these are retrospective reports. Knowledge is considered to be unconscious when subjects show an effect of training (e.g., above-chance performance on a classification task), despite being unable to describe the knowledge that underlies their performance. Several studies have provided evidence for a dissociation between task performance and verbalizable knowledge (e.g., Berry & Broadbent, 1984; Broadbent, 1977; Green & Hecht, 1992; Rebuschat & Williams, 2006, 2009, 2012; Williams, 2005), and there is little doubt that exposure can result in unconscious knowledge if lack of verbalization is used as a criterion for implicitness.

Another possibility is to compare subjects' performance on direct and indirect tests. The direct test is a measure that explicitly instructs subjects to make use of their knowledge. The task encourages subjects to access all relevant conscious knowledge to perform it (St. John & Shanks, 1997). The indirect test assesses subjects' performance without instructing them to use their acquired knowledge. Ideally, subjects should not even know that they are being tested when performing on the indirect task. Knowledge is assumed to be unconscious if an indirect test clearly indicates a learning effect, even though a direct test shows no evidence of learning. For example, Jiménez, Méndez, and Cleeremans (1996) used the Serial Reaction Time (SRT) task as an indirect test and a generation task as a direct test. Subjects performed on the two tasks successively. In the SRT task, subjects saw a stimulus appear at one of several locations on a computer screen and were asked to press, as fast and accurately as possible, on the corresponding key. Unbeknownst to subjects, the sequence of successive stimuli was determined by an artificial grammar. In the generation task, subjects were asked to predict the location of the next stimulus by pressing the corresponding key. Jiménez et al. (1996) found that subjects had clearly learned to exploit the regularities inherent in the stimulus environment, as evidenced by increased speed and accuracy of responding to grammatical sequences in the indirect SRT task, compared with ungrammatical sequences. Interestingly, whereas subjects were also able to express some knowledge in the direct (generation) task, further analyses indicated that large sequence learning effects were exclusively expressed through decreased response time in the SRT task. Their results suggest that at least some of the acquired knowledge was unconscious. Moreover, in implicit learning research, the view that knowledge can be unconscious when it is applied without conscious intention has gained popularity. Jacoby (1991) and Jacoby, Ste-Marie, and Toth (1993) introduce the Process Dissociation Procedure (PDP; see Curran [2001], Destrebecqz and Cleeremans [2001], and Wilkinson and Shanks [2004] for applications to implicit learning research). In the PDP paradigm, a participant performs an inclusion and an exclusion task. The inclusion part requires participants to apply acquired knowledge (and combines implicit and explicit processes), whereas the exclusion part requires them to suppress the acquired knowledge (which requires to control explicit processes while implicit processes interfere with the task). Accordingly, performance differences between both tasks provide some information regarding implicit and explicit processes or knowledge affecting the performance.

Finally, one can also determine whether subjects acquired conscious or unconscious knowledge by collecting confidence ratings and judgments about the kind of knowledge used to make the decision (e.g., Dienes & Scott, 2005; Dienes & Seth, 2010). This can be done, for example, by asking subjects to perform on a grammaticality-judgments task and to indicate, for each judgment, how confident they were in each decision (e.g., guess, somewhat confident, very confident) and what their decision was based on (e.g., guess, intuition, memory, rule knowledge). Subjects' knowledge can be considered unconscious when (a) the grammaticality judgments for which subjects reported to be guessing are actually significantly above chance (guessing criterion) and (b) when there is no correlation between the reported confidence level and the observed level of accuracy (zero correlation criterion) (see Dienes et al., 1995). In both of these cases, subjects are not aware of having acquired knowledge. There are several studies that have relied on subjective measures to provide evidence for unconscious knowledge (e.g., Dienes & Scott, 2005; Rebuschat & Williams, 2012; Tagarelli, Borges Mota, & Rebuschat, 2011).

3. Which musical features can be acquired?

To underpin the fundamental hypothesis that music cognition across cultures is grounded in (implicitly) acquired knowledge and to ask which musical features are acquired, implicit learnability has to be demonstrated for major types of structures and building blocks across cultures. This links implicit learning research closely to music theory and ethnomusicology (cf. Stevens, 2012). At present, most research was performed for melodic structure, yet a range of studies explored harmony, timbre, and rhythm (see Table 1).

3.1. Melody

Melody involves the temporal sequence of note events, representing at least pitch, onset, and duration (including rests) of a single (monophonic) voice. Melodic learning was frequently explored with respect to learning of statistical features, artificial grammars, and cross-serial dependencies.

It has been well established for a long time that adults and infants are able to pick up and apply tone statistics from musical exposure (cf. Castellano, Bharucha, & Krumhansl, 1984; Krumhansl, 1990; Krumhansl & Keil, 1982; Loui & Wessel, 2008; Loui, Wessel, & Hudson Kam, 2010). In analogy to language, adults as well as infants can learn a number of three-tone ''words'' from a continuous melodic stream built from concatenating these ''words'' (Saffran, Johnson, Aslin, & Newport, 1999; Saffran, Loman, & Robertson, 2000). Schön et al. (2008) further found that statistical learning of three-syllable words was enhanced

Table 1
Overview of musical features and structures explored by implicit or statistical learning

Feature	Description or Example	Structure	Learning Effect	Knowledge Type	(Selected) References
Timbre	Sequences of trumpet, gong, violin, etc. clap, drill	Finite-state grammar	✔	nt	Bigand et al. (1998) Howard and Ballas (1980, 1982)
	Timbre triplets (timbre words)	Chunks	✔	nt	Tillmann and McAdams (2004)
Melody	Probabilities of single pitches	Chunks (of size 1)	✔	nt	Castellano et al. (1984), Loui and Wessel (2008)
	Tone triplets (tone words)	Chunks	✔	nt	Saffran et al. (1999)
	AxByCz (e.g., BD#G#F#C#C)	Interleaved tone words	✗ *	nt	Creel et al. (2004)
	AxByzC (tonal & musically meaningful, e.g., $D_5A_4B_4A_4F\#_4$ G_4)	Interleaved tone words	✔	nt	Endress (2010)
	Whole melodies, testing for melody types or continuation tones	Exemplars, chunks	✔	nt	Rosner and Meyer (1982, 1986) Thompson et al. (2000) Szpunar et al. (2004)
	Tone sequences	Finite-state grammar	✔ Chunks	C	Rohrmeier et al. (2011)
			pitch probabilities	UC	Rohrmeier and Cross (2010)
				nt	Tillmann and Poulin-Charronnat (2010)
				nt	Loui and Wessel (2008)
				nt	Loui et al. (2010)
	$A_1B_1C_1D_1A_2B_2C_2D_2$ (cross-serial dependencies)	Context sensitive	✗ / ✔	UC	Dienes and Longuet-Higgins (2004)
			✔	UC	Kuhn and Dienes (2005)
			✗ / ✔ Chunks, grammar	C	Kuhn and Dienes (2006)
	Classicial North Indian rãgas (ecological melodic structure)	Chunks (melodic features)	✔	C	Rohrmeier & Widdess (2012)
Melody & sung syllables	Sung syllable sequences (tone-syllable-words)	Chunks	✔	nt	Schön et al. (2008) Clément and Schön (2010)
Harmony	Chord sequences	Finite-state grammar	✔	nt	Jonaitis and Saffran (2009)
				C	Bly et al. (2009)
	Chord sequences	Exemplar	✔	nt	Loui et al. (2009)

Table 1
Continued

Feature	Description or Example	Structure	Learning Effect	Knowledge Type	(Selected) References
	Chord sequences	Context-free grammar	Chunks, grammar	UC	Rohrmeier and Cross (2010), Rohrmeier (2010), ch. 4
Rhythm	Fixed duration patterns combined with another ordinal pattern (e.g., visual or pitch)	Exemplars, chunks	✔	UC	Buchner and Steffens (2001)
			Learning dependently & independently of corresponding ordinal pattern	UC/C	Shin and Ivry (2002)
				nt	O'Reilly et al. (2008)
				UC	Salidis (2001)
				UC	Ullén and Bengtsson (2003)
				UC / C	Karabanov and Ullén (2008)
	As above, but duration patterns with underlying Meter	Exemplars, chunks		nt	Tillmann et al. (2011)
				UC	Brandon et al. (2012)
				UC	Schultz et al. (in press)

Unexplored features:
Voice-leading, musical schemata, more interactions between features, prolongation structure, higher order structure / musical form, real ecologically valid (Non-Western) music.

Notes. The knowledge type represents more than one entry if the results from different studies differ. nt, not tested (statistical learning); C, conscious knowledge; UC, unconscious knowledge; ✔, success in learning; ✗, no success in learning.

*Learning only when tones were separated by streaming (timbre or two octaves apart).

when those were sung, and best when each of the syllables were matched with pitches. Employing a similar paradigm (but using EEG methodologies), Clément and Schön (2010) found that event-related potentials showed an implicit sensitivity to learned musical structure that could not be revealed by the behavioral test and, further, that musicians acquired both musical and linguistic structures better than nonmusicians (Clément & Schön, 2011). Omigie and Stewart (2011) found that statistical learning is fully intact in participants who suffered from congenital amusia, a disorder which affects pitch processing and production (Ayotte, Peretz, & Hyde, 2002; Peretz, Champod, & Hyde, 2003). However, Loui and Schlaug (2012) and Peretz, Saffran, Schön, and Gosselin (2012) found contrasting results of impaired statistical learning in amusics; hence, further research is required to settle these mixed results. When the tone words in a statistical-learning study were interleaved so that they were noncontiguous (e.g., AxByCz), participants could only learn the tone words when streaming was induced by pitch distance of two octaves or different timbres for both tone words (Creel, Newport, & Aslin, 2004). Endress (2010), however, found that participants were able to acquire nonadjacent tone dependencies in melodies when they were musically meaningful (in a major or minor key) instead of randomly constructed interleaved tone patterns (exp. 4).

Beyond these short patterns, several studies have explored the learning of larger melodic sequences organized by a finite-state grammar. Adults acquired short melodic structures generated by selecting sequences of single pitches from a short underlying harmonic sequence using an unfamiliar tuning based on the Bohlen–Pierce scale (Loui & Wessel, 2008; Loui et al., 2010)—stimuli that were isomorphic with a simple finite-state grammar (Loui & Wessel, 2008). Moreover, the exposure with melodies generated from a finite-state grammar affects the priming of target tones and the formation of melodic expectations (Tillmann & Poulin-Charronnat, 2010). Adults are further able to incidentally acquire novel diatonic, finite-state melodic structures that are substantially longer than the previously mentioned studies and feature up to 32 notes, and even acquire the knowledge of chunks online, that is, in real time, during a test (Rohrmeier, Rebuschat, & Cross, 2010), suggesting a powerful learning mechanism. However, when the melodies generated from the same grammar intentionally violated Narmour's cross-cultural melodic principles (Krumhansl, 1995; Narmour, 1990), the learning performance was reduced (Rohrmeier & Cross, 2010), which suggests that implicit learning is facilitated by established common structures in the system (e.g., music; see also Loui, 2012; Chen et al., 2011).

Few studies explored the learning of ecologically valid entire melodies or melodic structures. Rohrmeier and Widdess (2012) employed a cross-cultural paradigm to investigate learning of North Indian music. An entire Alap section (an ''exposition''part) from two different modes (rāgas) was employed for the learning phase for two groups. The testing phase used different melodic fragments from a later section of the identical performance that exhibited typical features of either rāga. Results suggest that even after short exposure, participants had acquired some, but not all (Non-Western) features from both rāgas. Because confidence ratings were correlated with performance, findings suggested that participants possessed explicit judgment knowledge. Using entire tonal ecological target melodies, Rosner and Meyer (1982, 1986) found that participants could learn the melodic ''gap-fill'' archetype (Meyer, 1956, 1973; although see von Hippel, 2000; for methodological criticisms). Thompson, Balkwill, and Vernescu (2000) (experiment 3) tested melodic continuations of the final melody note after exposure to entire melodies. Their results suggest exemplar or chunk learning from mere exposure (awareness was not tested). A follow-up study by Szpunar, Schellenberg, and Pliner (2004) found that participants liked melodies that they were exposed to better than (even very similar) novel melodies.

All findings above could largely be explained by chunk learning (cf. Rohrmeier & Koelsch, 2012). In particular, the acquisition of tone profiles corresponds learning the probabilities of single pitch classes (chunks of length 1; ''unigrams''). Similarly, implicit knowledge of Narmour's (1990) melodic principles can be embraced by learning of melodic interval pairs (''bigrams''; cf. Pearce & Wiggins, 2006). Learning of long sequences or whole melodies may equally be explained by the acquisition of large chunks.

Beyond musical structures that could be explained by chunks, a different series of studies explored implicit learning of more complex, abstract rule-based melodic transformations. In this paradigm, melodic structures are constructed which consist of two halves, the second of which is a transform of the first half, for example, $A_1B_1C_1D_1$ $A_2B_2C_2D_2$, in which X_1 and X_2

correspond by a specific transformation such as inversion (mutually replacing upwards or downwards intervals) or retrograde (having the order of pitches reversed) or a combination of both (inverse-retrograde). In general, such structures are also frequently referred to as ''cross-serial dependencies'' or ''bigrammatical'' structures. Dienes and Longuet-Higgins (2004) explored learning of cross-serial structures under a serialist paradigm using dodecaphonic (i.e., twelve tone) stimuli and inverse-retrograde and inversion transformations. They found that only selected expert participants could implicitly learn to detect stimulus melodies that exhibited the above serialist structures. Similarly, Kuhn and Dienes (2005) found that participants could acquire implicit knowledge of cross-serial dependencies using diatonic inversions controlling for knowledge of chunks and stimulus exemplars. In another study, however, Kuhn and Dienes (2006) compared incidental and intentional learning conditions and found that only participants who were encouraged to search for a rule were able to learn the inversion pattern (similar to findings outside the musical domain, cf. Johnstone & Shanks, 2001; Mathews et al., 1989; Shanks et al., 1997). Cross-serial structures cannot be captured by the mere acquisition of fragments or *n*-grams; thus, the set of results is still debated. Desmet, Poulin-Charronnat, Lalitte, and Perruchet (2009) criticized that the findings by Kuhn and Dienes (2005) might be flawed due to a confound that grammatical melodies contained a higher probability of an ambitus of an entire octave than ungrammatical melodies. However, Dienes and Kuhn (unpublished) showed that the effect remained significant, statistically controlling for the confound (as well as another retrograde confound). Further it was later shown that a similar inversion in poetry (Jiang et al 2012) and movement (Dienes et al, 2012) could be learned, when the confounds discussed by Desmet et al had been controlled.

3.2. Harmony

Harmony refers to the sequence of chords that are built from a set of at least two concurrent pitches and conceptualized as if sounding simultaneously and as having a specific root. For instance, the pitches {C3, E3, A3} may be represented as an A-minor chord with the root A and C as the bass note. Jonaitis and Saffran (2009) found that two groups of participants acquired statistical information about chord transitions generated from either of two finite-state grammars and tested from examples of both grammars. The materials were built from major, minor, and diminished chords within the context of the less common Phrygian scale. A comparable effect of statistical learning of chord frequencies, based on a short chord progression and a deviant chord in a non-Western tuning system, was found by Loui, Wu, Wessel, and Knight (2009). Similarly, another study by Bly, Carrión, and Rasch (2009) compared implicit learning of artificial, finite-state harmony sequences that followed or violated some basic principles of diatonic Western harmony and found learning in both conditions; however, the performance for the finite-state grammar violating harmony principles was significantly lower, matching the findings about melodic structure by Rohrmeier and Cross (2010). The control group performance, which underpinned that prior knowledge of the tonal system enhanced the learning of the stimuli, was not sufficient to explain the learning performance of the regular or irregular stimulus sequences. Although finite-state grammars constitute a major paradigm for the exploration of implicit knowledge, there is theoretical evidence that harmony is rather governed by more complex, context-free dependencies (Rohrmeier, 2007, 2011; Steedman, 1984, 1996; Tojo, Oka, & Nishida, 2006). A study with musician participants found them to acquire a novel harmonic system under the PDP paradigm. The harmonic system was constructed on the octatonic scale and based on a context-free grammar (Rohrmeier & Cross, 2009). In this study, grammatical structure was the strongest predictor, whereas bigram learning was not a predictor or, at best, only a weak predictor (exp. 2). They further found that materials with greater complexity (structures generated from center-embedding recursion rather than tail recursion) were less well learned than the respective simpler sequences.

3.3. Timbre and rhythm

The artificial grammar-learning paradigm was also employed for timbre sequences of musical timbres (e.g., trumpet, gong, violin) (Bigand, Perruchet, & Boyer, 1998), or environmental sounds (such as hand clap or drill) (Howard & Ballas, 1980, 1982). Tillmann and McAdams (2004) expanded the statistical-learning paradigm by Saffran et al. (1999) and found that participants could acquire the segmentation of a continuous stream of musical timbres. None of these studies, however, tested awareness of the learning and thus it is not granted that the learned knowledge is in fact implicit.

Although some studies examine implicit learning of temporal patterns, only few studies addressed the learning of musical rhythm and meter. The evidence concerning implicit learning of temporal patterns is ambiguous. Three SRT studies found participants unable to learn temporal patterns independently of other (correlated) ordinal visual or pitch structures (Buchner & Steffens, 2001; O'Reilly, McCarthy, Capizzi, & Nobre, 2008; Shin & Ivry, 2002). However, there are several contrasting results: Salidis (2001) adapted the SRT paradigm (cf. Nissen & Bullemer, 1987) for the learning of rhythmical structures based on beeps of the same pitch (using response-stimulus intervals in contrast with stimulus inter-onset intervals). Beeps in three different temporal intervals (180, 450, 1250 ms) were played either in a specific or random order and participants were required to press a key as quickly as possible after a beep. Findings indicated a decrease in reaction time and implicit learning of the temporal sequence. Using the PDP paradigm, Ullén and Bengtsson (2003) found independent implicit learning of audiovisual ordinal and temporal patterns. Karabanov and Ullén (2008) presented participants with random audiovisual sequences combined with a specific temporal pattern. In a serial recall task, they instructed participants to reproduce either the ordinal sequence or both the temporal and ordinal sequences, and found evidence for implicit learning of temporal sequence when no instruction to reproduce the temporal pattern was given. Adapting the SRT paradigm, Brandon, Terry, Stevens, and Tillmann (2012) similarly found incidental learning of temporal sequence of duple or triple metrical structure combined with uncorrelated pseudo-random sequences of three syllables (using exposure inter-onset interval [IOI] patterns of 13111122 and 12311112, based on a temporal unit of 700 ms). Participants showed no explicit knowledge of metrical temporal structures.

Several of the above studies, for example, Salidis (2001), Shin and Ivry (2002), or O'Reilly et al. (2008), employed temporal patterns that were not representative of typical musical rhythms. Adapting the SRT framework, Tillmann, Stevens, and Keller (2011) employed a simple, specifically musical short–short–long rhythm of two synthesized syllables and tones using durations of 2–2–3 beat units compared with a simpler 2–2–4 rhythm. The former implies a 7/8 metre (of a cycle of seven quaver beats), which is rare in common Western music (in contrast to, e.g., Bulgarian music, see London, 2004), whereas the latter implies a frequent rhythmic pattern in a 4/4 metre. For both rhythms, they found that exposure resulted in decreased response times and improved tapping task performance and temporal expectations. Because the tempi of the test and training stimuli differed, they concluded that participants learned relative rather than absolute timing patterns.

Schultz, Stevens, Keller, and Tillmann (in press) employed a modified PDP framework to investigate learning of temporal patterns with strong, weak, or no metricality (following Povel & Essens, 1985) using a single pitch sound that was pseudo-randomly presented through the left, right, or both channels. Results showed implicit learning of the temporal structure without a correlated ordinal pattern in a single- but not multiple-response SRT task. They argued that probabilistic uncertainty of an upcoming ordinal stimulus identity may explain the discrepancy (above) of why some previous studies (Miyawaki, 2006; O'Reilly et al., 2008; Shin, 2008; Shin & Ivry, 2002) had found no independent learning of temporal and ordinal patterns. Schultz et al. further examined a hypothesis drawn from the Dynamic Attending Theory (Jones & Boltz, 1989), predicting that strongly metrical rhythms would be more readily learned than nonmetrical rhythms (with complex IOI ratios that are not attributable to a metrical framework). However, temporal expectancies were acquired for both metrical and nonmetrical patterns and, in contrast with the hypothesis, nonmetrical patterns were not less readily learned. Nonetheless, within the metrical condition, expectations were stronger for events occurring in strong compared with weak metrical locations, suggesting an effect of metric binding (Jones, 2009).

Given the complexity of rhythmical and metrical structure in music and its rich interaction with pitch, melody, and harmony, the present results provide a basis concerning temporal learning; however, more research is required for a comprehensive understanding of the learning of musical temporal regularities and the integration of temporal and pitch features.

3.4. Discussion: Which types of musical structures are learned?

The presented studies provide evidence for different musical features, but they also relate to the general debate concerning the types of structures that can be acquired implicitly. As argued, diverse results, including the learning of tone profiles, local but not nonlocal tone words, pitch or chord transition probabilities, events beginning or ending a sequence (anchor positions) or implied continuations of intervals, could be embraced by fragment learning. Evidence for context-free structures or cross-serial dependencies, however, challenge mere fragment-learning hypothesis, but further research is required to settle the debate. Accepting these findings, there are no overarching (computational) learning models predicting all these bits of evidence (Rohrmeier, 2010).

4. Computational models of (implicit) musical learning: Findings and open issues

Many experimental studies draw assumptions or sketch models regarding characteristics of the learning mechanism or the mental representation of the acquired knowledge. Wiggins (2011) argues that such experimentally based models may be called ''pen-and-paper models,'' and that they rather constitute ''verbal'' computational models that are specified in greater or lesser detail and not (yet) implemented. For instance, the assumption that implicit learning corresponds largely or solely to the learning of chunks, element statistics, or transition probabilities (e.g., Perruchet & Pacton, 2006) constitutes such a hypothesized (and frequently underspecified, see below) model that can be implemented and evaluated in precise computational ways (see also, e.g., models such as *n*-gram models, Competitive Chunker, or PARSER, discussed below). Correspondingly, computational modeling forces all overt and covert assumptions to be made explicit and therefore affords the precise implementation and evaluation of hypothetical processes employing reference studies or critical empirical data; predictions of computational models can be examined under conditions that are impossible to perform experimentally (such as long-term exposure). There are many computational models of music composition, music information retrieval, or specific features of music perception or prediction (cf. Dubnov, 2010; Rohrmeier & Koelsch, 2012; Wang, 2010). Most of these models are learning models, because models with static assumptions were frequently found incapable of dealing the complexity of real-world music (cf. Wiggins, 2011; see recent ISMIR conference proceedings)—which constitutes evidence against the innateness of complex mechanisms of musical processing. Despite the plenitude of models, not much cognitive work has yet been done bridging behavioral and computational approaches to implicit learning of music. In this review, we select models specifically with a focus to this cognitive question.

Because music notation represents music using a small, discrete alphabet of pitches, temporal intervals, and harmony (and key structure), models operating on such a symbolic representation were facilitated and frequently applied with success (cf. Wiggins, 2011). Generally, modeling approaches to music can largely be divided into symbolic and neural-network models reflecting modeling different levels of abstraction and explanation (Cleeremans & Dienes, 2008; Marr, 1982, 1982; Newell, 1973; Toiviainen, 2000; Weizenbaum, 1976). While symbolic fragment-based models assume that chunk learning suffices to explain various features of music perception, neural networks are set to model connectivity principles of underlying neural structures and show that features of tonal organization can be picked up by self-organization and connectionist learning (see Table 2 for an overview). The underlying learning algorithms are further distinguished according to whether the learning is supervised or unsupervised.[1] The implicit learning that people perform inside or outside the laboratory corresponds to unsupervised learning from largely positive examples (during acquisition, people mostly do not hear incorrect structures based on which they could adjust inferred overgeneralizations).

Table 2
Overview of main, selected cognitive computational models of (implicit) learning of music, arranged by the types of features and types of models

Feature	Input Representation	Model	Objective	Type of Training	References (Selection)
Signal	Signal	Self-Organizing network	Timbre similarity	Unsupervised	Toiviainen et al. (1995) Toiviainen (1996) Toiviainen et al. (1998a,b)
Melody	Symbolic (pitch)	*n*-gram	Prediction generation method comparison	Unsupervised	Pearce and Wiggins (2004)
	Symbolic feature streams (e.g., pitch, duration, scale degree, pitch* duration, etc.)	Feature-based and combined *n*-gram models (multiple-viewpoint model) IDyOM	Prediction, generation	Unsupervised	Conklin and Witten (1995) Pearce (2005) Pearce and Wiggins (2006)
	Symbolic	Self-Organizing network	Recognition, prediction, generation	Unsupervised	Louhivuori et al. (1996) Krumhansl et al. (1999) Krumhansl et al. (2000)
	Symbolic (pitch)	Simple Recurrent Network Feed-forward network	Learning cross-serial dependencies	Supervised	Kuhn and Dienes (2008)
Harmony	Symbolic (single chords and streams of features)	*n*-gram Hidden Markov Models Dynamic Bayesian Networks	Prediction, generation	Unsupervised	Whorley et al. (2010) Rohrmeier and Graepel (2012)
Key tonality	Symbolic (pitches)	Self-Organizing network	Key inference dynamics	Unsupervised	Toiviainen and Krumhansl (2003)
Tonality	Symbolic (pitches, chords, keys)	MUSACT static activation network (predefined connections between features)	Priming and interaction of pitch, harmony, & key	Supervised	Bharucha (1987)
	Symbolic	Self-Organizing network (prefigured network layers)	Feature interaction and tonality (pitch, harmony, & key)	Unsupervised	Tillmann et al. (2000)
Discrete event sequences	Symbolic (single, discrete event stream)	Competitive Chunker PARSER	Fragment learning	Unsupervised	Servan-Schreiber and Anderson (1990) Perruchet and Vinter (1998)

4.1. Symbolic and probabilistic models

Fragment models implement chunking hypotheses that learning is based on extracting, storing, and combining small contiguous chunks independently of their neural implementation. They are grounded on the condition that their input stream is symbolic and discrete (rather than signal based or continuous). Various *n*-gram models, Information Dynamics of Music (IDyOM), Competitive Chunker, and PARSER ground on the fragment hypothesis idea (Pearce, 2005; Perruchet & Vinter, 1998; Servan-Schreiber & Anderson, 1990). Most simply, unsupervised *n*-gram models learn information from sequences by chopping them into short fragments (*n*-grams) up to a size of n and compute predictions based on the statistics of the accounted *n*-grams using different methods for recombining, and smoothing (accounting for yet unobserved cases); these issues are mostly ignored in pen-and-paper models of chunking (e.g., Perruchet & Pacteau, 1990; for a comparison of different methods see Pearce & Wiggins, 2004). As argued above and elsewhere (Rohrmeier & Koelsch, 2012), *n*-gram models are sufficiently general to embrace theories of chunk learning, statistical learning of transition frequencies, as well as whole exemplar learning (very long *n*-grams). While *n*-gram models are restricted to one single stream of events, the multiple viewpoint idea (Conklin & Witten, 1995) expands *n*-gram methods to modeling the interaction of information of parallel streams of low-level features (e.g., pitch, duration, onset) and (precoded) high-level musical features (e.g., scale degree, scale degree combined with duration, onset combined with duration, etc.); it constitutes the heart of the IDyOM model (Pearce & Wiggins, 2012). Multiple-viewpoint models use dependencies and redundancy between different musical features to enhance its predictive power (search techniques like genetic algorithms are applied to approach optimal feature combinations, Pearce, 2005).

From another perspective, Competitive Chunker (Servan-Schreiber & Anderson, 1990) and PARSER (Perruchet & Vinter, 1998) exceed mere *n*-gram models with higher order integration of smaller chunks into hierarchically or sequentially organized larger chunks, but they have not yet extensively been applied to music. In a model comparison of a range of implicit learning experiments, *n*-gram models were found very powerful and to largely outperform humans, whereas Competitive Chunker was close to the human range, and PARSER below it (Rohrmeier, 2010). Without presuming innate principles of melodic processing, the simple mechanism of IDyOM has been shown to explain various behavioral results of melodic prediction and segmentation as well as electroencephalographic (EEG) data for melodic prediction (Pearce & Wiggins, 2006; Pearce, Müllensiefen, & Wiggins, 2010a; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010b; see Pearce & Wiggins, 2012 for more details). They further predict statistical online learning (Pearce et al., 2010b). While most of these models were predominantly applied to melodic prediction or segmentation, only preliminary implementations for harmony exist (Ponsford et al., 1999; Rohrmeier, 2007b; Rohrmeier & Graepel, 2012; Whorley, Wiggins, Rhodes, & Pearce, 2010).

Although models other than *n*-gram models for discrete sequences are frequently used in music information retrieval, they are rarely applied for cognitive modeling. Hidden Markov Models (HMM), a graphical generalization of Markov models that predict surface symbols with different distributions based on the probabilistic changes of a hidden state, were successful for music analysis or generation (e.g., Allan & Williams, 2005; Raphael & Stoddard, 2004). One difference between HMM models and *n*-gram models is that the former can take longer sequential context more flexibly into account than the latter. In analogy to the multiple-viewpoint technique, Dynamic Bayesian Networks (DBN) and input–output models (Murphy, 2002) constitute a generalization of HMMs that allows the learning of feature combinations. Rohrmeier and Graepel (2012) implemented such a model for harmony prediction and found that it outperformed plain HMM models and slightly outperformed multiple-viewpoint *n*-gram models in predicting chords based on past context and mode information. Using a similar graphical model, Paiement et al. (2009) proposed a specific type of input–output HMMs for the modeling of melodic prediction based on harmonic structure.

4.2. Neural networks

On a different modeling level, neural network models have been frequently used for modeling different features of music production, music information retrieval, or perception (Bharucha, 1987; Griffith & Todd, 1999; Leman, 1997). In general, three types of models were explored: activation networks, feed-forward or simple recurrent networks (SRN, Elman, 1990), and self-organizing maps (SOM, Kohonen, 1995). Early connectionist models of music (Bharucha, 1987) were successful in predicting some features of tonal perception (see Bharucha, 2011, for an overview), yet they presumed the levels of representation (tone, chord, key) in an activation network in a hard-wired way: Pitches were mapped to chords and those were mapped to keys according to music theoretical standards, which renders the success of the network to predict musical experimental results less surprising (see Wiggins, 2011).

Several studies employed Self-Organizing Maps (SOM; Kohonen, 1995), which are capable of unsupervised learning, modeling timbre similarity (Toiviainen, 1996; Toiviainen, Kaipainen, & Louhivuori, 1995; Toiviainen et al., 1998), or melodic recognition, expectancy, and generation (Krumhansl, Louhivuori, Toiviainen, Järvinen, & Eerola, 1999; Krumhansl et al., 2000; Louhivuori, Kaipainen, & Toiviainen, 1996). A SOM model of the dynamics of tonality perception, modeling the relationships between different keys (Toiviainen & Krumhansl, 2003), was found to correlate with the dynamic topography of the rostromedial prefrontal cortex (Janata et al., 2002). Tillmann, Bharucha, and Bigand (2000) built a three-layer Self-Organizing Map based on pitch input. After targeted training with artificial pitch configurations (forming chords) or short sequences, they found the SOM to converge toward theoretically meaningful organizations of pitch, chord, and key structure. Subsequent analysis of the model matched some experimental reference results of tonal cognition, such as harmonic priming (however, Wiggins [2011] pointed out that more complex issues in the inference of tonal and harmonic knowledge such as segmentation or boundary detection remained unaddressed).[2]

In contrast to fragment and HMM models, one challenge in the connectionist modeling of music is to capture the temporality of music as a sequence of different pitches and durations. Stevens and Latimer (1997a,b) chose a static representation for modeling melody recognition. Their perceptron model featured one input layer that coded different static predefined time slices by different single input neurons (coding pitch by a real number between 0 and 1) that were connected to two output neurons coding the two possible recognition responses. Their network matched behavioral results using different melodies of identical predefined length well; however, the restricted static representation has limitations in capturing long real-world melodies and potentially capturing temporal shifts in pitch sequences. There are different modeling techniques that capture the temporality of a sequence without using an unrolled representation of activation. Following up the work by Gjerdingen (1989, 1990), Page (1994) created a variant (employing masking fields, Cohen & Grossberg, 1987) of a self-organizing adaptive resonance network (Carpenter & Grossberg, 1987), which formed realistic patterns of melodic expectancy from a small training set of 12 nursery tunes. Elman's (1990) Simple Recurrent Networks (SRNs) endow a simple three-layer feed-forward network with an additional set of hidden-context units which store the past activation of the hidden layer and connect to the activation of the present hidden layer. This way, information about activation patterns from past elements in a sequence can be stored. Mozer (1991, 1994) built an SRN for melodic generation and prediction based on a multiple representation (cf. Shepard, 1982) of pitch using pitch height, chroma, cycle of fifths, and relative duration. The network outperformed bigram models but did not generalize to overarching melodic structure.

Buffer models (Boucher & Dienes, 2003) model temporality using a sliding window similar to *n*-gram techniques: Their architecture is a simple three-layer feed-forward network, for which the input layer codes the past *n* events simultaneously. During training or evaluation of a sequence, the sliding input window is shifted through the sequence. SRNs were found to learn finite-state grammars and transition frequencies (cf. Cleeremans, 1993; Cleeremans & Dienes, 2008). Kuhn and Dienes (2008) showed that they were further able to represent nonlocal cross-serial dependencies (modeling the behavioral results by Kuhn & Dienes, 2005, 2006), without requiring the less elegant buffer model (Boucher & Dienes, 2003).

Cleeremans and Dienes (2008) as well as Honing (2006), stress that modeling specific human characteristics, or generating surprising predictions, is more informative than tweaking models to an optimal fit. In this spirit, Kuhn and Dienes (2008) found that SRN and buffer models fell into the characteristic human performance range of Kuhn and Dienes (2005). Contrariwise, none of blank-slate *n*-gram, Chunker, or SRN models could capture the characteristic decay of human performance when finite-state melodies violated Narmour's principles (discussed above; Rohrmeier, 2010; Rohrmeier & Cross, 2010). When the models were initially pretrained with a folksong corpus, *n*-gram and SRN models exhibited some decay in performance which, however, did not fully match the human pattern. The best-fitting Chunker model, in contrast, did not show any difference between the Narmour-consistent and Narmour-inconsistent set of stimuli. Hence, mere pretraining does not suffice as a hypothesis to explain the experimental findings, and additional preprocessing or streaming mechanisms may be assumed.

There is a fundamental difference between the learning studies discussed above and the computational models: Most experimental approaches focus specifically on the musical features and types of structures that can be acquired, and their representation, which is tested typically by specific yet typically ecologically unrealistic, classification or recognition tasks using whole sequences (although see Tillmann & Poulin-Charronnat, 2010 or Loui et al., 2009; for innovative procedures). Computational models, however, employ specific representations and learning algorithms to model particular musical perceptual tasks like prediction, segmentation, or recognition or to provide predictors for neuroscience techniques (EEG or fMRI). The learning task is regarded successful when the learning algorithm suffices to match the perceptual task. The question as to which (minimal) representation is required to capture specific formal structures and how different forms of knowledge representation differ for specific tasks is, however, not much addressed. Furthermore, although many studies evaluate computational models against experimental data, little is known about how well several different models perform across a number of different sets of data. In a Bayesian model comparison, Rohrmeier (2010, ch. 8) found that most *n*-gram models outperformed humans, whereas Competitive Chunker and SRN models generally matched human performance characteristics best.

Although there is a range of successful cognitive learning models of music, much research remains to be done. At present there are no computational models which unify the different levels of musical structure (ranging from signal [timbre or spectral], symbolic note level, or higher order syntactic or formal structures), structural complexity, or perceptual tasks. Further research is required to better understand the interactions between temporal and structural features, as well as the learning of non-Western musical structures which do not employ discrete elements. Finally, the difference between implicit and explicit knowledge is not addressed at all, as well as the ways in which implicit knowledge can become explicit (cf. Cleeremans & Dienes, 2008). Finally, most models concentrate on modeling learning and perception of a single individual, whereas the ways in which social interaction gates learning are scarcely modeled (cf. Cross, 2012; Sun, Coward, & Zenzen, 2005).

5. General discussion

By now there is a body of research addressing learning of music from experimental and computational perspectives. Yet, do the results fully support that music cognition across cultures is grounded in (implicitly) acquired knowledge?

Although implicit learning research is targeted at musical acquisition, enculturation, or stylistic learning, many of the studies on melodic structure and pitch explore uncharacteristic stimuli and sidelines of music perception: Timbre ‘‘words’’ or grammars are not generally prototypical for real-world music, similarly melodic structure, for instance, does not frequently employ ‘‘tone words,’’ isochronous sequences, strict cross-serial dependencies, or biconditional grammars; melody can, however, like any set of sequences be approximated by large, probabilistic finite-state grammars or large chunk collections. Although the use of such abstract and frequently simple reference structures constitutes an initial straightforward step taken in many domains, the obtained findings subsequently only provide weak support for the implicit acquisition of real-world musical structure. The difference between simplicity and parsimony of learning mechanisms and representations conflicts with theoretical assumptions about musical structure: From experimental and computational perspectives, learning of statistical tones distributions and transitions seems plausible; many of such assumptions have, however, risen from past choices of initial simple experimental paradigms, stimulus structures, or simple algorithms. Music theory in contrast, stressed the complex interaction between temporal and pitch features, and the hierarchical, context-free (tree-structured) nature of melody, voice leading, or harmony (Lerdahl & Jackendoff, 1983; Rohrmeier, 2011; Schenker, 1935). Although the assumption of simple learning mechanisms is ubiquitous, the benchmark for success should be adequate real-world Western and non-Western musical structures (cf. contributions of Cross, 2012; Stevens, 2012). Much research is necessary, employing realistic, ecologically valid stimuli and addressing common theoretically underpinned building blocks, for example, regarding melodic, voice-leading or harmonic schemata (cf. Byros, 2009; Gjerdingen, 1988, 2007), formal (Caplin, 1998), modal, prolongational (Lerdahl & Jackendoff, 1983; Rohrmeier, 2011) patterns and feature interactions. While cross-cultural perspectives stress the fundamentally intersubjective and interactive nature of music (Cross, 2003; Keller, 2008), there are (yet) no results (nor research paradigms) exploring implicit learning in interactive musical contexts, and the extent to which social interaction (between musicians or audience/participants) gates or aids implicit learning.

Most studies report results with respect to structures that we can learn, while only center-embedding (context-free) structures or cross-serial dependencies are debated. Because structures we are unable to acquire provide fundamental insights ruling out particular mechanisms of learning or representation, more experimental research, particularly negative results (with respect to the structures that participants cannot learn implicitly), and modeling is required to explore the limits and boundaries of our structural learning capabilities.

Besides musical acquisition and enculturation, another important aspect of musical listening concerns picking up individual patterns of a piece during listening, that is, learning in real time (online or real-time learning). The learning of such patterns (such as tone profiles or chunks) relates to the same process as in long-term acquisition. *n*-gram models provide predictions about online learning (Pearce & Wiggins, 2012) that were in part experimentally tested (Potter, Wiggins, & Pearce, 2007). Similarly, Rohrmeier et al. (2010) found that a group of untrained participants acquired melodic chunk knowledge during the test phase of an artificial grammar-learning experiment. This finding corresponds with a series of above-chance control group results in implicit learning experiments as well as common real-time learning results in SRT tasks (Clegg, DiGirolamo, & Keele, 1998). It indicates a strong learning process which could be predicted by three computational models (Rohrmeier, 2009; Rohrmeier, 2010). These findings of online learning open various avenues for future research regarding new methodologies, the learning speed, and the interaction or interference of such knowledge with long-term knowledge.

Altogether, the study of implicit learning and knowledge in music provides substantial contributions in the general cognitive science domain. Music contains structures that are comparably complex as language (Patel, 2008; Pearce & Rohrmeier, 2012; Rebuschat, Rohrmeier, Cross, & Hawkins, 2011) and uses an enormous range of cognitive functions (Alluri et al., 2012). It can be studied without needing to worry about the interaction with propositional semantics and conceptual world knowledge. Musical learning and modeling research provide insights and challenges into core human cognitive abilities that inform debates (cf. Pearce & Rohrmeier, 2012), for example, about the learning of context-free or cross-serial dependencies, recursive or nonrecursive processing, multiple-stream processing, feature integration and attention, and interactions across domains such as the facilitation of language learning by musical structure (such as by singing). Such contributions render music research one core branch of the cognitive sciences.

Notes

1. *Supervised learning* refers to a learning task for which feedback is available during the training phase, that is, training data are presented as labeled training data (as correct or incorrect, or belonging to other classes) or input and desired output combinations (as frequently used for feed-forward neural network models). *Unsupervised learning* refers to learning conditions when no such information is available in training materials (i.e., there is no reward or error feedback that the learning algorithm can use).
2. Wiggins further argues that the SOMs were originally designed as tools of scientific visualization inspired by neural activity (Kohonen, 1995), and that hence cognitive implications, (expected and unexpected) predictions about cognitive behavior and the level of modeling require clear specification.

References

Allen, R., & Reber, A. S. (1980). Very long term memory for tacit knowledge. *Cognition*, *8*, 175–185.

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage*. *59*(4), 3677–3689.

Ayotte, J., Peretz, I., & Hyde, K. L. (2002). Congenital Amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238–251.

Berry, D. C. (1997). *How implicit is implicit learning?* Oxford, England: Oxford University Press.

Berry, D. C., & Broadbent, D. (1984). On the relationship between task performance and associated verbalisable knowledge. *Quarterly Journal of Experimental Psychology*, *36*, 209–231.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hove, England: Lawrence Erlbaum.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, *5*, 1–30.

Bigand, E., Perruchet, P., & Boyer, M. (1998). Implicit learning of an artificial grammar of musical timbres. *Cahiers de Psychologie Cognitive-Current Psychology of Cognition*, *17*(3), 577–600.

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.

Bly, B. M., Carrión, R. E., & Rasch, B. (2009). Domain-specific learning of grammatical structure in musical and phonological sequences. *Memory & Cognition*, *37*(1), 10–20.

Boucher, L., & Dienes, Z. (2003). Two ways of learning associations. *Cognitive Science*, *27*, 807–842.

Brandon, M., Terry, J., Stevens, C. J., & Tillmann, B. (2012). Incidental learning of temporal structures conforming to a metrical framework. *Frontiers in Psychology*, *3*, 294. doi: 10.3389/fpsyg.2012.00294

Broadbent, D. (1977). Levels, hierarchies and the locus of control. *Quarterly Journal of Experimental Psychology*, *29*, 181–201.

Buchner, A., & Steffens, M. C. (2001). Simultaneous learning of different regularities in sequence learning tasks: Limits and characteristics. *Psychological Research*, *65*, 71–80.

Byros, V. (2009). Towards an ''archaeology'' of hearing: Schemata and eighteenth-century consciousness. *Musica Humana*, *1*(2), 235–306.

Caplin, W. (1998) *Classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven*. New York: Oxford University Press.

Carpenter, G. A., & Grossberg, S. (1987). RT 2: Self- organization of stable category recognition codes for analog input patterns. *Applied Optics*, *26*, 4919–4930.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, *113*, 394–412.

Chan, C. (1992). *Implicit cognitive processes: Theoretical issues and applications in computer systems design*. Unpublished PhD thesis, Oxford, England: University of Oxford.

Chen, W., Guo, X., Tang, J., Zhu, L., Yang, Z., & Dienes, Z. (2011). Unconscious structural knowledge of form-meaning connections. *Consciousness and Cognition*, *20*, 1751–1760.

Chomsky, N. (1957). *Syntactic structures*. The Hague, The Netherlands: Mouton.

Cleeremans, A. (1993). *Mechanisms of implicit learning: Connectionist models of sequence processing*. Cambridge, MA: MIT Press.

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

Cleeremans, A., & Dienes, Z. (2008). Computational models of implicit learning. In R. Sun (Ed.), *Cambridge handbook of computational psychology* (pp. 396–421). Cambridge, England: Cambridge University Press.

Clegg, B. A., DiGirolamo, G. J., & Keele, S. W. (1998). Sequence learning. *Trends in Cognitive Sciences*, *2*, 275–281.

Clément, F., & Schön, D. (2010). Learning of musical and linguistic structures: Comparing event-related potentials and behavior. *NeuroReport*, *21*(14), 928–932.

Clément, F., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357–2365.

Cohen, M. A., & Grossberg, S. (1987). Masking fields: A massively parallel neural architecture for learning, recognizing, and predicting multiple groupings of patterned data. *Applied Optics*, *26*, 1866–1891.

Conklin, D., & Witten, I. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, *24*(1), 51–73.

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.

Cross, I. (2003). Music and biocultural evolution. In M. Clayton, T. Herbert, & R. Middleton (Eds.), *The cultural study of music: A critical introduction* (pp. 19–30). New York: Routledge.

Cross, I. (2012). Cognitive science and the cultural nature of music. *Topics in Cognitive Science*. *4*, 668–677.

Curran, T. (2001). Implicit learning revealed by the method of opposition. *Trends in Cognitive Sciences*, *5*, 503–504.

Dalla Bella, S., Peretz, I., & Aronoff, N. (2003). Time course of melody recognition: A gating paradigm study. *Perception and Psychophysics*, *65*(7), 1019–1028.

Deliège, I. (2001). Prototype effects in music listening: An empirical approach to the notion of imprint. *Music Perception*, *18*(3), 371–407.

Deliège, I., Mèlen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real time listening to a piece of music. *Music Perception*, *14*(2), 117–160.

Deliège, I., & Sloboda, J. A. (1995). *Musical beginnings: Origins and development of musical competence*. Oxford, England: Oxford University Press.

Desmet, C., Poulin-Charronnat, B., Lalitte, P., & Perruchet, P. (2009). Implicit learning of nonlocal musical rules: A comment on Kuhn and Dienes (2005). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 299–305.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*(2), 343–350.

Dienes, Z. (2009). Artificial grammar learning. In P. Wilken, T. Bayne & A. Cleeremans, (Eds.), *Oxford companion to consciousness* (pp. 585–588). Oxford, England: Oxford University Press.

Dienes, Z. (2011). Conscious versus unconscious learning of structure. In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 337–364). Berlin, Germany: Mouton de Gruyter Publishers.

Dienes, Z., & Altmann, G. (1997). Transfer of implicit knowledge across domains? How implicit and how abstract? In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 107–123). Oxford, England: Oxford University Press.

Dienes, Z., & Berry, D. C. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin and Review*, *4*, 3–23.

Dienes, Z., Broadbent, D. E., & Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 875–882.

Dienes, Z., Kuhn, G., Guo, X. Y., & Jones, C. (2012). Communicating structure, affect and movement: Commentary on Bharucha, Curtis & Paroo. In Rebuschat, P., Rohrmeier, M., Cross, I. Hawkins (Eds.), *Language and Music as Cognitive Systems. Oxford University Press* (pp. 156–168).

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. Structural vs judgement knowledge. Psychological Research. *Psychological Research*, *69*, 338–351.

Dienes, Z., & Seth, A. K (2010). The conscious and the unconscious. In G. Koob, R. F. Thompson & M. Le Moal (Eds.), *Encyclopedia of behavioral neuroscience* (vol. 1, pp. 322–327). Oxford, England: Academic Press.

Dubnov, S. (2010). Musical information dynamics as models of auditory anticipation. In W. Wang (Ed.), *Machine audition: Principles, algorithms and systems* (pp. 371–397). Hershey, PA: IGI Global Publication.

Dulany, D. E., Carlson, R., & Dewey, G. (1984). A case of syntactical learning and judgement: How concrete and how abstract? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *113*, 541–555.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211.

Endress, A. D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica*, *135*(2), 182–190.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251.

Ettlinger, M., Margulis, E. H., & Wong, P. C. M. (2011). Implicit memory in music and language. *Frontiers in Psychology*, *211*(2), doi: 10.3389/fpsyg.2011.00211

Francès, R. (1958/1988). *The perception of music*. In W. J. Dowling (Trans.). Hillsdale, NJ: Erlbaum.

Fu, Q. F., Fu, X. L., & Dienes, Z. (2008). Implicit sequence learning and conscious awareness. *Consciousness and Cognition*, *17*, 185–202.

Gaillard, V., Vandenberghe, M., Destrebecqz, A., & Cleeremans, A. (2006). First- and third-person approaches in implicit learning research. *Consciousness and Cognition*, *15*, 709–722.

Gjerdingen, R. O. (1988). *A classic turn of phrase: Music and the psychology of convention*. Philadelphia, PA: University of Pennsylvania Press.

Gjerdingen, R. (1989). Using connectionist models to explore complex musical patterns. *Computer Music Journal*, *13*, 67–75.

Gjerdingen, R. (1990). Categorization of musical patterns by self-organizing Neuronlike networks. *Music Perception*, *7*, 339–370.

Gjerdingen, R. O. (2007). *Music in the galant style*. Oxford, England: Oxford University Press.

Gómez, R. L. (2007). Statistical learning in infant language development. In M. Gareth Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 601–616). Oxford, England: Oxford University Press.

Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics*, *13*, 168–184.

Greenwald, A. G. (1992). New look 3: Unconscious cognition reclaimed. *American Psychologist*, *47*, 766–779.

Griffith, N., & Todd, P. M. (Eds.) (1999). *Musical networks: Parallel distributed perception and performance*. Cambridge, MA: MIT Press.

Hallam, S., Cross, I., & Thaut, M. (Eds.) (2008). *Oxford handbook of music psychology*. Oxford, England: Oxford University Press.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*, 466–472.

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

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

von Hippel, P. (2000). Questioning a melodic archetype: Do listeners use gap-fill to classify melodies? *Music Perception*, *18*(2), 139–153.

Honing, H. (2006). Computational modeling of music cognition: A case study on model selection. *Music Perception*, *23*, 365–376.

Honing, H. (2009). *Iedereen is muzikaal. Wat we weten over het luisteren naar muziek*. Amsterdam: Nieuw Amsterdam Uitgevers.

Howard, J. H. J., & Ballas, J. A. (1980). Syntactic and semantic factors in the classification of nonspeech transient patterns. *Perception & Psychophysics*, *28*, 431–439.

Howard, J. H. J., & Ballas, J. A. (1982). Acquisition of acoustic pattern categories by exemplar observation. *Organization, Behavior and Human Performance*, *30*, 157–173.

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

Huron, D. (2012). Two challenges in cognitive musicology. *Topics in Cognitive Science*, *4*, 678–684.

Jackendoff, R. (1991). Musical parsing and musical affect. *Music Perception*, *9*, 199–230.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*, 33–72.

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

Jacoby, L. L., Ste-Marie, D., & Toth, J. P. (1993). Redefining automaticity: Unconscious influences, awareness, and control. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control* (pp. 261–282). Oxford, England: Oxford University Press.

Janata, P., Birk, J. L., Van Horn, J.D., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, *298*, 2167–2170.

Jiang, S., Zhu, L., Guo, X., Ma, W., Yang, Z., & Dienes, Z. (2012). Unconscious structural knowledge of tonal symmetry: Tang poetry redefines limits of implicit learning. *Consciousness & Cognition*, *21*(1), 476–486.

Jiménez, L., Méndez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 948–969.

Johnstone, T., & Shanks, D. R. (2001). Abstractionist and processing accounts of implicit learning. *Cognitive Psychology*, *42*, 61–112.

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

Jones, M. R. (2009). Musical time. In S. Hallam, I. Cross & M. Thaut (Eds.), *The handbook of music psychology* (pp. 81–92). New York: Oxford University Press.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491.

Karabanov, A., & Ullén, F. (2008). Implicit and explicit learning of temporal sequences studied with the process dissociation procedure. *Journal of Neurophysiology*, *100*, 733–739.

Keller, P. E. (2008). Enacting intersubjectivity: A cognitive and social perspective to the study of interactions. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective to the study of interactions* (pp. 205–221). Amsterdam: IOS Press.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of category-level knowledge and explicit memory for specific instances. *Psychological Science*, *3*, 172–179.

Koelsch, S. (2011). Toward a neural basis of music perception—a review and updated model. *Front Psychology*, *2*, 110. doi: 10.3389/fpsyg.2011.00110

Koelsch, S., Gunter, T., & Friederici, A. (2000). Brain indices of music processing: ''Nonmusicians'' are musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541.

Kohonen, T. (1995). *Self-organizing maps*. Berlin: Springer-Verlag.

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

Krumhansl, C. L. (1995). Effects of musical context on similarity and expectancy. *Systematische Musikwissenschaft*, *3*, 211–250.

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

Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception*, *17*(2), 151–196.

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition*, *75*, 1–46.

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

Kuhn, G., & Dienes, Z. (2006). Differences in the types of musical regularity learnt in incidental- and intentional-learning conditions. *The Quaterly Journal of Experimental Psychology*, *59*(10), 1725–1744.

Kuhn, G., & Dienes, Z. (2008). Learning non-local dependencies. *Cognition*, *1*, 184–206.

Leman, M. (Ed.), (1997). *Music, gestalt, and computing. Studies in cognitive and systematic musicology*. Berlin and Heidelberg: Springer.

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

Lewicki, P., Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, *20*, 24–37.

Louhivuori, J., Kaipainen, M., & Toiviainen, P. (1996). A self-organizing map that recognizes and generates rune-song melodies. *Systematische Musikwissenschaft*, *IV/1-2*, 149–164.

Loui, P. (2012). Learning and liking of melody and harmony: Further studies in artificial grammar learning. *Topics in Cognitive Science*, *4*, 554–567.

Loui, P., & Schlaug, G. (2012). Impaired learning of event frequencies in tone-deafness. *Annals of the New York Academy of Sciences*, *1252*(1), 354–360.

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

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

Loui, P., Wu, E. H., Wessel, D. L., & Knight, R. T. (2009). A generalized mechanism for perception of pitch patterns. *Journal of Neuroscience*, *29*(2), 454–459.

Marcus, G. (2012). Musicality: Instinct or acquired skill? *Topics in Cognitive Science*, *4*, 499–513.

Marr, D. (1982). *Vision: A computational approach*. San Francisco: Freeman & Co.

Matthews, R., Buss, R., Stanley, W., Blancard-Fields, F., Cho, J-R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1083–1100.

Merikle, P. M., & Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Consciousness and Cognition*, *6*, 219–236.

Meyer, L. B. (1956). *Emotion and meaning in music*. London: University of Chicago Press.

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

Miyawaki, K. (2006). The influence of the response-stimulus interval on implicit and explicit learning of stimulus sequence. *Psychological Research*, *70*, 262–272.

Mozer, M. C. (1991) Connectionist music composition based on melodic, stylistic and psychophysical constraints. In P. M. Todd & D. G. Loy (Eds.), *Music and connectionism* (pp. 195–212). Cambridge, MA: MIT Press.

Mozer, M. C. (1994). Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multi-scale processing. *Connection Science*, *6*(2–3), 247–280.

Murphy, K. (2002). *Dynamic Bayesian networks: Representation, inference and learning*. Unpublished Ph.D. dissertation, University of California at Berkeley.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.

Newell, A (1973). Artificial intelligence and the concept of mind. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language* (pp. 1–60). San Francisco: W. H. Freeman.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.

Omigie, D., & Stewart, L. (2011) Preserved statistical learning of tonal and linguistic material in congenital Amusia. *Frontiers in Psychology*, *2*(109). doi: 10.3389/fpsyg.2011.00109

O'Reilly, J. X., McCarthy, K. J., Capizzi, M., & Nobre, A. C. (2008). Acquisition of the temporal and ordinal structure of movement sequences in incidental learning. *Journal of Neurophysiology*, *99*, 2731–2735.

Page, M. P. A. (1994). Modelling the perception of musical sequences with self-organizing neural networks. *Connection Science*, *6*(2–3), 223–246.

Paiement, J-F., Bengio, S., & Eck, D. (2009). Probabilistic models for melodic prediction. *Artificial Intelligence*, *173* (14), 1266–1274.

Patel, A. D. (2008). *Music, Language, and the Brain*. Oxford, England: Oxford University Press.

Payne, J. W. (1994). Thinking aloud: Insights into information processing. *Psychological Science*, *5*, 241–248.

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition*. London: City University.

Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2010a). The role of expectation and probabilistic learning in auditory boundary perception: A model comparison. *Perception*, *39*, 1367–1391.

Pearce, M. T., & Rohrmeier, M. A. (2012). Music and the cognitive sciences. *Topics in Cognitive Science*, *4*, 468–484.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010b). Unsupervised statistical learning underpins computational, behavioural and neural manifestations of musical expectation. *NeuroImage*, *50*, 302–313.

Pearce, M. T., & Wiggins, G. A. (2004). Improved methods for statistical modelling of monophonic music. *Journal of New Music Research*, *33*(4), 367–385.

Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception*, *23*(5), 377–405.

Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, *4*, 625–652.

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of musical disorders. The Montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences*, *999*, 58–75.

Peretz, I., Saffran, J., Schön, D., & Gosselin, N. (2012). Statistical learning of speech, not music, in congenital amusia. *Annals of the New York Academy of Sciences*, *1252*, 361–367.

Perruchet, P. (2008). Implicit learning. In H. Roediger (ed.), *Cognitive psychology of memory. Vol. 2 of learning and memory: A comprehensive reference* (pp. 597–621). Oxford, England: Elsevier.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge. *Journal of Experimental Psychology: General*, *119*(3), 264–275.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238.

Perruchet, P., & Vinter, A. (1998). PARSER: A model of word segmentation. *Journal of Memory and Language*, *39*, 246–263.

Ponsford, D., Wiggins, G., & Mellish, C. (1999). Statistical learning of harmonic movement. *Journal of New Music Research*, *28*, 150–177.

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

Potter, K., Wiggins, G. A., & Pearce, M. T. (2007). Towards greater objectivity in music theory: Information-dynamic analysis of minimalist music. *Musicae Scientiae*, *11*(2), 295–322.

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

Raphael, C., & Stoddard, J. (2004). Functional harmonic analysis using probabilistic models. *Computer Music Journal*, *28* (3), 45–52.

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

Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, *81*, 115–119.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235.

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

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 888–896.

Rebuschat, P., Rohrmeier, M., Cross, I., & Hawkins, J. (Eds.) (2011) *Language and music as cognitive systems*. Oxford, England: Oxford University Press.

Rebuschat, P., & Williams, J. N. (2006). Dissociating implicit and explicit learning of natural language syntax. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (p. 2594). Mahwah, NJ: Lawrence Erlbaum.

Rebuschat, P., & Williams, J. N. (2009) Implicit learning of word order. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 425–430). Austin, TX: Cognitive Science Society.

Rebuschat, P., & Williams, J. N. (Eds.) (2012) *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, *33*, 829–856.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology*, *125*(2), 123–137.

Ritchie, W. C., & Bhatia, T. K. (Eds.) (2009). *New handbook of second language acquisition*. Bingley, England: Emerald Group Publishing Ltd.

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

Rohrmeier, M. (2007b) Modelling dynamics of key induction in harmony progressions. In H. Spyridis, A. Georgaki, G. Kouroupetroglou & C. Anagnostopoulou (Eds.) *Proceedings of the 4th Sound and Music Computing Conference*, pp. 82–89

Rohrmeier, M. (2009). Learning on the fly. Computational analyses of an unsupervised online-learning effect in artificial grammar learning. *Proceedings of the 9th International Conference on Cognitive Modelling*, Manchester, England.

Rohrmeier, M. (2010). *Implicit learning of musical structure: Experimental and compuational modelling approaches*. Unpublished PhD Thesis. Cambridge, England: University of Cambridge.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, *5*(1), 35–53.

Rohrmeier, M., & Cross, I. (2009). Tacit tonality: Implicit learning of context-free harmonic structure. In Louhivuori et al. (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music* (pp. 443–452). Jyväskylä, Finland: University of Jyväskylä.

Rohrmeier, M., & Cross, I. (2010). Narmour's principles affect implicit learning of melody. In S. Demorest et al. (ed.), *Proceedings of the 11th International Conference on Music Perception and Cognition (ICMPC 2010)*.

Rohrmeier, M., Fu, Q., & Dienes, Z. (in press). Implicit learning of recursive context-free grammars. *PLoS ONE*.

Rohrmeier, M., & Graepel, T. (2012). Comparing feature-based models of harmony. In M. Barthet et al. (Ed.) *Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR 2012)* (pp. 357–370). London: Queen Mary, University of London.

Rohrmeier, M., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, *83*, 164–175. DOI: 10.1016/j.ijpsycho.2011.12.010.

Rohrmeier, M., Rebuschat, P., & Cross, I. (2011). Incidental and online learning of melodic structure. *Consciousness and Cognition*, *20*(2), 214–222. doi:10.1016/j.concog.2010.07.004

Rohrmeier, M., & Widdess, R. (2012). Incidental learning of modal features of North Indian music. In E. Cambouropoulos, C. Tsougras, P. Mavromatis & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition (ICMPC 2012) and 8th Conference of the European Society for the Cognitive Sciences of Music (ESCOM 2012)* (pp. 857–866).

Rosner, B. S., & Meyer, L. B. (1982). Melodic processes and the perception of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 317–341). New York: Academic Press.

Rosner, B. S., & Meyer, L. B. (1986). The perceptual roles of melodic process, contour and form. *Music Perception*, *4*(1), 1–40.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *8*, 101–105.

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

Saffran, J. R., Loman, M., & Robertson, R. (2000). Infant memory for musical experiences. *Cognition*, *77*, B15–B23.

Salidis, J. (2001). Nonconscious temporal cognition: Learning rhythms implicitly. *Memory and Cognition*, *29*, 1111–1119.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, *106*, 975–983.

Schultz, B. G., Stevens, C. J., Keller, P. E., & Tillmann, B. (in press). The implicit learning of metrical and non-metrical temporal patterns. *Quarterly Journal of Experimental Psychology*.

Servan-Schreiber, D., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 592–608.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*(8), 314–321.

Shanks, D. R. (2005). Implicit learning. In Lamberts K. and Goldstone R.. editors. *Handbook of Cognition*, p. 202–220. Sage, London.

Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, *89*, 305–333.

Shin, J. C. (2008). The procedural learning of action order is independent of temporal learning. *Psychological Research*, *72*, 376–386.

Shin, J. C., & Ivry, R. B. (2002). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology Learning, Memory and Cognition*, *28*, 445–457.

St. John, M. F., & Shanks, D. R. (1997). Implicit learning from an information-processing standpoint. In D. C. Berry (Ed.), *How implicit is implicit learning?* (pp. 162–194). Oxford, England: Oxford University Press.

Stadler, M. A., & Roediger, H. L. (1998). The question of awareness in research on implicit learning. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 105–132). Thousand Oaks, CA: Sage Publications.

Stalinski, S., & Schellenberg, E. G. (2012). Music cognition: A developmental perspective. *Topics in Cognitive Science*, *4*, 485–497.

Steedman, M. J. (1984). A generative grammar for Jazz chord sequences. *Music Perception*, *2*(1), 52–77.

Steedman, M. J. (1996). The blues and the abstract truth: Music and mental models. In A. Garnham & J. Oakhill (Eds.), *Mental models in cognitive science* (pp. 305–318). Mahwah, NJ: Erlbaum.

Stevens, C. J. (2012). Music perception and cognition: A review of recent cross-cultural research. *Topics in Cognitive Science*, *4*, 653–667.

Stevens, C., & Latimer, C. (1997a). Music recognition: An illustrative application of a connectionist model. *Psychology of Music*, *25*, 161–185.

Stevens, C., & Latimer, C. (1997b). A comparison of connectionist models of music recognition and human performance. *Minds and Machines*, *2*, 379–400.

Sun, R., Coward, L. A., & Zenzen, M. J. (2005). On levels of cognitive modeling. *Philosophical Psychology*, *18*(5), 613–637.

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*(2), 370–381.

Tagarelli, K., Borges Mota, M., & Rebuschat, P. (2011). The role of working memory in implicit and explicit language learning. In L. Carlson, C. Hölscher & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2061–2066). Austin, TX: Cognitive Science Society.

Temperley, D. (2001). *The cognition of basic musical structures*, Cambridge, MA: MIT Press.

Thompson, W. F., Balkwill, L. L., & Vernescu, R. (2000). Expectancies generated by recent exposure to melodic sequences. *Memory and Cognition*, *28*(4), 547–555.

Tillmann, B. (2005). Implicit investigations of tonal knowledge in nonmusician listeners. *Annals of the New York Academy of Science*, *1060*, 100–110.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*(4), 885–913.

Tillmann, B., & McAdams, S. (2004). Implicit learning of musical timbre sequences: Statistical regularities confronted with acoustical (dis)similarities. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *30*(5), 1131–1142.

Tillmann, B., & Poulin-Charronnat, B. (2010). Auditory expectations for newly acquired structures. *Quarterly Journal of Experimental Psychology*, *63*(8), 1646–1664.

Tillmann, B., Stevens, C., & Keller, P. E. (2011). Learning of timing patterns and the development of temporal expectations. *Psychological Research*, *75*(3), 243–258.

Toiviainen, P. (1996). Optimizing auditory images and distance metrics for self-organizing timbre maps. *Journal of New Music Research*, *25*(1), 1–30.

Toiviainen, P. (2000). Symbolic AI versus connectionism in music research. In E. Miranda (Ed.), *Readings in music and artificial intelligence* (pp. 47–68). Amsterdam: Harwood Academic Publishers.

Toiviainen, P., Kaipainen, M., & Louhivuori, J. (1995). Musical timbre: Similarity ratings correlate with computational feature space distances. *Journal of New Music Research*, *24*(3), 282–298.

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception*, *32*(6), 741–766.

Toiviainen, P., Tervaniemi, M., Louhivuori, J., Saher, M., Huotilainen, M., & Näätänen, R. (1998a). Timbre similarity: Convergence of neural, behavioral, and computational approaches. *Music Perception*, *16*(2), 223–241.

Tojo, S., Oka, Y., & Nishida, M. (2006). Analysis of chord progression by HPSG. In (Ed.), *Proceedings of the 24th IASTED International Conference on Artificial Intelligence and Applications* (pp. 305–310). Innsbruck, Austria.

Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. Riess-Jones & R. R. Fay (Eds.), *Springer handbook of auditory research: Music perception* (pp. 89–128). Heidelberg, Germany: Springer.

Trehub, S. E. (2006). Infants as musical connoisseurs. In G. McPherson (Ed.), *The child as musician: A handbook of musical development* (pp. 33–49). Oxford, England: Oxford University Press.

Trehub, S., & Hannon, E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, *100*, 73–99.

Ullén, F., & Bengtsson, S. L. (2003). Independent processing of the temporal and ordinal structure of movement sequences. *Journal of Neurophysiology*, *90*, 3725–3735.

Wang, W. (Ed.) (2010) *Machine audition: Principles, algorithms and systems*. Hershey, PA: IGI Global Publication.

Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. San Francisco: W H Freeman.

Whorley, R., Wiggins, G. A., Rhodes, C. S., & Pearce, M. T. (2010). Development of techniques for the computational modelling of harmony. In D. Ventura, A. Pease, R. Pérez y Pérez, G. Ritchie & T. Veale (Eds.), *Proceedings of the International Conference on Computational Creativity* (pp. 11–15). Lisbon, Portugal: Department of Informatics Engineering, University of Coimbra.

Wiggins, G. (2011). Computational models of music. In P. Rebuschat, M. Rohrmeier, I. Cross & J. Hawkins (Eds.), *Music and language as cognitive systems*. Oxford, England: Oxford University Press.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 354–369.

Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, *27*(2), 269–304.

Williams, J. (2009). Implicit learning. In W. C. Ritchie & T. K. Bhatia (Ed.), *New handbook of second language acquisition* (pp. 319–353). Bingley, England: Emerald Group Publishing Ltd.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 2.

Zajonc, R. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, *10*(6), 224–228.