Music Cognition: Learning, Perception, Expectations

Barbara Tillmann

University Claude Bernard Lyon 1, Sensory Neurosciences, Behavior and Cognition Laboratory, CNRS-UMR 5020, IFR 19,
50, avenue Tony Garnier, F-69366 Lyon cedex 07, France
btillmann@olfac.univ-lyon1.fr
http://olfac.univ-lyon1.fr/equipe-02.html

Abstract. Research in music cognition domain has shown that non musician listeners have implicit knowledge about the Western tonal musical system. This knowledge, acquired by mere exposure to music in everyday life, influences perception of musical structures and allows developing expectations for future incoming events. Musical expectations play a role for musical expressivity and influence event processing: Expected events are processed faster and more accurately than less-expected events and this influence extends to the processing of simultaneously presented visual information. Studying implicit learning of auditory material in the laboratory allows us to further understand this cognitive capacity (i.e., at the origin of tonal acculturation) and its potential application to the learning of new musical systems and new musical expectations. In addition to behavioral studies on cognitive processes in and around music perception, computational models allow simulating learning, representation and perception of music for non musician listeners.

1 Immersion in Music: What Is the Brain Doing?

In everyday life, we are immersed almost constantly in a musical environment. The development of mp3-players and music-playing telephones has been further enhancing this immersion. In Western culture, most of this music (e.g., classical music, pop, rock, folk, jazz, lullabies) is based on the Western tonal system, even if new musical styles as well as music of other cultures are increasingly present. Most of music listeners are without explicit musical training or practice on an instrument: how is the nonmusician brain process and understand musical structures? The present chapter will focus on perceivers, not composers or performers (see [1, 2] for reviews on musical performance).

Music cognition research has provided evidence that nonmusician listeners have acquired implicit knowledge about the Western tonal musical system, just by mere exposure to musical pieces obeying the rules of this system. The musical knowledge, acquired thanks to the cognitive capacity of implicit learning, influences the perception of musical structures and allows developing expectations for future incoming events. This chapter presents some of the basic regularities of the tonal system and an

R. Kronland-Martinet, S. Ystad, and K. Jensen (Eds.): CMMR 2007, LNCS 4969, pp. 11–33, 2008.

overview of empirical data showing that listeners have knowledge about these regularities and that this knowledge influences music perception (section 2). Particular emphasize is given to studies investigating listeners' musical expectations and their influence on event processing, both auditory and visual, as well as their link to musical expressivity (section 3). Section 4 focuses on the cognitive capacity of implicit learning, which is the basis of tonal acculturation, and how it can be studied in the laboratory with new artificial tone and timbre systems. Section 5 presents how connectionist models can be used to simulate nonmusician listeners, notably for the learning of a musical system, the cognitive representation of this knowledge and its influence on perception. Finally, most music cognition research has studied learning, perception and expectations for music of the Western tonal system. However, the same questions apply to the processing of other musical systems (section 6). To show the generality of the cognitive capacity of learning, knowledge and expectations it is necessary that the research domain overcomes the Western tonal focus and shows comparable phenomenon for other musical systems (see [3]).

2 Tonal Knowledge and Perception of Musical Structures

The overall pattern of results in music cognition research suggests that mere exposure to Western musical pieces suffices to develop implicit knowledge of the tonal system. Just by listening to music in everyday life, listeners become sensitive to the regularities and structures of the tonal system without being necessarily able to verbalize them [4-6]. This acquisition is based on the cognitive capacity of implicit learning (see section 5). The implicitly acquired knowledge influences listeners' music perception, the understanding of musical structures and relations as well as the development of musical expectations (see section 3).

The present section proposes a summary of the Western tonal system underlining its statistical regularities (i.e., forming the basis of musical structures and relations) and reviews some experimental research investigating listeners' tonal knowledge and its influence on perception.

2.1 Some Basic Regularities in Western Tonal Music

Western tonal music can be described as a constrained system of regularities (i.e., regularities of co-occurrence, frequency of occurrence and psychoacoustic regularities) based on a limited number of elements. This section presents the tonal system from the perspective of cognitive psychology and of implicit learning: it underlines the basic regularities between musical events, which appear in most musical styles of occidental everyday life (e.g., classical music, pop music, jazz music, Latin music etc.) and which can be acquired by implicit learning processes.[1]

The Western tonal system is based on 12 pitches repeated cyclically over octaves. Strong regularities of co-occurrence and frequencies of occurrence exist among the 12

[1] In addition to these regularities based on the pitch dimension, regularities exist on the time dimension, like the underlying beat allowing listeners to develop temporal expectations about when the next event is the most probable to occur. This chapter focuses on the pitch dimension, but will consider the time dimension in sections 3 and 6.

pitch classes: tones are combined into chords and into keys, forming a three-level organizational system. Sets of 7 tones form scales, which can be either major or minor. For each tone of a scale, chords (e.g., major or minor) are constructed by adding two tones – creating a second level of musical units. Based on tones and chords, keys (tonalities) define a third level of musical units. Keys have more or less close harmonic relations to each other, with the strength of harmonic relations depending on the number of shared tones and chords. The three levels of musical units (i.e., tones, chords, keys) occur with strong regularities of co-occurrence. Tones and chords belonging to the same key are more likely to co-occur in a musical piece than tones and chords belonging to different keys. Changes between keys are more likely to occur between closely related keys than between less-related ones.

Within each key, tones and chords have different tonal functions creating tonal and harmonic hierarchies. For example, chords built on the first, fifth, and fourth scale degrees (referred to as tonic, dominant and subdominant respectively) have a more central function than chords built on other scale degrees. From a psychological point of view, the hierarchically important events of a key act as stable cognitive reference points [7] to which other events are anchored [8]. These within-key hierarchies are strongly correlated with the frequency of occurrence of tones and chords in Western musical pieces. Tones and chords used with higher frequency (and longer duration) correspond to events that are defined by music theory as having more important functions in a given key [5, 6, 9].

This short description reveals a fundamental characteristic of Western tonal music: the tonal functions of tones and chords are depending on the established key context; the same event can define an in-key or an out-of-key event and can take different levels of functional importance (i.e., tonal stability). For example, the *C* major chord functions as a stable tonic chord in a *C* major context, as a less stable dominant or subdominant chord in *F* or *G* major contexts respectively and as an out-of-key chord in a *D* major context. For listeners, understanding this context dependency of musical events' functions is crucial for the understanding of musical structures, notably the variety of musical structures that can be created on the basis of the restricted set of twelve pitch classes. Listeners' tonal knowledge is necessary to interpret differently the same sound event depending on the context in which it occurs. Acculturated listeners understand these musical structures in an implicit way.

2.2 Listeners' Tonal Knowledge of Pitch Structures

Numerous research has provided evidence for nonmusicians' tonal knowledge of pitch structures (see [5, 10, 11] for reviews). Nonmusician listeners are sensitive to the context dependency of musical events: they perceive the musical structures and relations between tones, chords and keys (i.e., modulations) (e.g., [12-14]). This knowledge also allows the perception of the underlying harmonic structures as well as the understanding of musical motifs and their variations. Behavioral studies in music cognition domain have provided evidence for this tonal knowledge with a variety of experimental methods. The seminal probe-tone paradigm, which asks listeners to rate how well a given tone fits into a preceding tonal context, showed the perceived context dependency for tones as well as the perceived distances between keys (see [6] for a review). For other subjective judgments, listeners rate the musical material for the degree of perceived similarity, musical tension or completion (e.g., [15, 16]). These judgments of perceived musical tension or completion reflect tonal hierarchy, as described by music theory: stable tones and chords (i.e., with important tonal functions) receive lower ratings of musical tension and were judged to induce more completion than less stable tones and chords. The influence of the perceived pitch structures has been shown also in memory tasks (e.g., [17]) and speeded response time tasks, which will be presented more specifically in section 3.

Experimental research on music perception has to find a balance between the complexity of real musical material and strong experimental control of the used material – leading to the use of simple tone contexts or melodies. One attempt to push the balance in the direction of real material used short minuets to investigate the perception of tonal functions [18]. A minuet is often constructed in two parts, with a first part ending on a half cadence (i.e., with the dominant, a less stable degree) and the second part on an authentic cadence (i.e., with the tonic, the most stable degree). Completion judgments on these parts presented separately showed that listeners perceive the differences in tonal stability of these endings: parts ending on the authentic cadence received higher ratings of completion than did parts ending on the half cadence. Furthermore, results of a musical jigsaw puzzle (i.e., participants had to find, among other choices, the two parts belonging to the same piece and to put them in the correct temporal order) showed that even nonmusician listeners understand these structural markers and their role for the temporal organization of the musical piece.

This overview presents some examples about how we can experimentally access listeners' musical knowledge and study its influence on perception. Tonal knowledge works like a "perceptual filter" and it allows listeners to understand the context dependency of musical events. Listeners interpret musical events with their knowledge about basic regularities of the tonal system, and they develop musical expectations for future events depending on the context. Because of the temporal nature of sound, expectations are taking a central position in music perception: which sounds are most probable to come next and when should they occur?

3 Musical Expectations: Pitch and Time

When presented with a musical context (i.e., the beginning of a melody), listeners develop expectations about future events – what kind of event (tone, chord, timbre) is the most probable to occur next and at what time point. These musical expectations can be linked to sensory features (or surface features), such as dynamic and timbral characteristics, melodic contour and intervals, but also to the repetition of specific notes, note groups or motifs. They can be further linked to the tonal structures of the musical system and thus require listeners' tonal knowledge, acquired by mere exposure to musical pieces obeying this system.

Expectations are part of cognitive functions in general, they are shaping our interaction with the environment and (depending on their nature) facilitate or delay the processing of events. The processing of an expected event is facilitated in comparison to an unexpected or less-expected event. For example, we recognize a familiar face faster in a context, in which it is probable to appear and is thus expected (i.e., our neighbor in front of his house) than in an unrelated context (i.e., on vacation in the mountains). Musical expectations are not only influencing efficiency of processing, but have been attributed a role for musical expressivity. Composers (or improvising performers) fulfill listeners' perceptual expectations more or less early or only partially in the musical piece. Following Meyer [19], playing with musical expectations is at the origin of musical expressivity that is communicated by musical pieces. In the following, we first present experimental research interested in the efficiency of processing for musically expected events and this research line brings us back to musical expressivity and emotion.

Musical expectations have been studied with production and perception tasks. In production tasks, participants hear a musical context (i.e., two tones or the first bars of a musical piece) and were requested to produce the continuation. In Carlsen [20, 21], participants were asked to sing what they felt to be the most natural continuation. In Schmuckler [22, 23], pianists were asked to complete fragments in terms of how they expected the piece would continue. In perception tasks, expectations are investigated by either asking participants to directly judge the "expectedness" of a musical phrase's ending on a subjective scale (from 1 unexpected to 7 expected; [23, 24]) or measuring speed of processing (with the hypothesis that expected events are processed faster than unexpected ones). Production tasks are mainly limited to the investigation of musical experts, while perception tasks allow the investigation of nonmusician listeners. Of particular interest among perception tasks is the priming paradigm, an indirect investigation method of perceiver's contextual expectations. The present section reviews this paradigm and its application to the investigation of tonal and temporal expectations, cross-modal influences as well as schematic *versus* veridical expectations.

3.1 The Priming Paradigm: Studying Nonmusicians' Musical Expectations

The priming paradigm (extensively used in psycholinguistics, see [25]) is an implicit investigation method that studies the influence of perceivers' expectations on the efficiency of perception (i.e., accuracy and processing speed). This implicit investigation method allows probing nonmusicians' musical knowledge without requiring explicit judgments (see [26] for a review).

In this paradigm, a prime context (i.e., a chord sequence) is followed by a target event (i.e., a chord) and the relation between prime and target is systematically manipulated (i.e., musical relatedness as defined by music theory). The hypothesis is that the prime context allows listeners to develop expectations for future events, with more strongly related events being more expected. These expectations then influence event processing, notably processing is facilitated for expected events over unexpected or less-expected events.

Since the priming paradigm is an indirect investigation of the context's influence on event processing, participants are not required to make direct judgments on the relation between prime context and target, but their task focuses on another dimension of the target event. Participants make speeded judgments on a perceptual feature of the target; a frequently used task is based on sensory consonance/dissonance judgments (e.g., [27, 28]), and for this purpose, half of the targets are consonant (i.e., well-tuned, correctly constructed chords), half of the targets are rendered acoustically dissonant (i.e., either by mistuning or by adding out-of-key tones)[2]. In most studies, the manipulated relations between prime context and target event concerned pitch relations for chords (harmonic structures) and tones, but less often temporal relations.

Harmony and Melody. The manipulated pitch relations between prime context and target chord can be relatively strong, contrasting an expected, in-key target (i.e., the tonic) to an unexpected, out-of-key target [26, 33], or more subtle, comparing two chords belonging to the context key. For example, Bigand and collaborators used eight-chord sequences and the last chord defined the target. The target chord acted as either the most important chord of the context key and was supposed to be highly expected (the tonic chord) or a less important chord, supposed to be less expected (the subdominant chord). In order to reduce sensory influences, the target chord's relation to the global context (chords 1 to 6) was manipulated while holding constant the local context (chord 7). The requested consonant/dissonant judgments are more accurate and faster when targets act as tonic rather than as subdominant chords [28, 34]. This outcome, valid for both musician and nonmusician participants, suggests that the processing of harmonic spectra is facilitated for events that are the most predictable in the current key context. Global musical priming effects have been extended to longer contexts (14-chord sequences, [34]) and more severe control of sensory influences [35]. Furthermore, the processing advantage is not restricted to the comparison of tonic and subdominant chords, but processing times reflect the top of the tonal hierarchy: the tonic is processed the fastest, followed by the dominant and then the subdominant [36].

More recently, the musical priming paradigm, which was initially introduced solely for chord processing, was extended to melodic processing: melodies were constructed by pair and differed only by a single note so that the target functioned as either the tonic or the subdominant. Processing times were faster for the related tonic target tone. Using melodies and target tones (instead of target chords) further allowed us to investigate whether musical expectations influence perceptual processes (e.g., detection, pitch processing). In the described melodic material, melodic expectations (based on listeners' tonal knowledge) influenced pitch discrimination, with finer discrimination for the expected tonic tones [37].

Time. Although pitch is the most obvious form-bearing dimension of Western tonal music, regularities in other musical dimensions also contribute to listeners' perceptual experience and may be internalized through similar processes. Beyond pitch, time is a crucial form-bearing dimension in music [38]. Temporal regularities include the organization of event-onset-intervals through time leading to a sensation of meter - a sensation of a regular succession of strong and weak beats superimposed over an isochronous pulse. Temporal regularities also include the temporal patterns of onset intervals creating rhythms that are perceived against the metrical background.

---

[2] When the experimental manipulations contrast related and unrelated target chords, the experimental trials consist of musical sequences with 25% of the trials ending on related consonant chords, 25% on related dissonant chords, 25% on unrelated consonant chords and 25% on unrelated dissonant chords. Additional priming tasks require judgments of temporal asynchrony [29], phoneme-discrimination [30], timbre-discrimination [31] and lexical decision of sung words/nonwords [32].

Temporal regularities have been shown to influence the perception of musical events in many ways, including performance in recognition tasks [28, 39-41], recall [42], completion judgments [40, 43, 44], evaluations of musical tension [15, 45], and musical expectations [24, 34, 40, 46, 47]. To investigate the influence of temporal expectations on chord processing, the priming paradigm was adapted to manipulate temporal structures, notably by opposing regular, isochronous sequences to irregular sequences and by manipulating the temporal occurrence of the last chord [40]. These manipulations are similar to those previously used with subjective judgments [24, 43]. In the priming paradigm, processing was facilitated when sequences were played in a regular, isochronous way in comparison to when played irregularly. In addition, processing was slowed down when targets occurred earlier than expected in comparison to on-time or later than expected.

For music perception, the question is how listeners process pitch and time dimensions together since their combination defines the musical structure of a piece (e.g., [48]). For the respective contributions of tonal and temporal regularities in music processing, two theoretical frameworks have been distinguished [41]. A single-component model [46] predicts interactive interference between the processing of the two dimensions. A two-component model, based on experimental and neuropsychological data, predicts that tonal and temporal structures are processed independently, and the processing of one dimension does not interfere with the processing of the other dimension. The hypothesis, which is currently proposed in music cognition domain, is that independence between the two dimensions occurs at initial stages of processing, but is followed by integration of the two dimensions in later stages of processing, thus leading to interactive influences [40, 41, 49, 50].

3.2 Cross-Modal Influences

Musical expectations based on listeners' tonal knowledge are fast and automatic (see also section 3.3). Their influence is not restricted to the processing of musical features, but influences the processing of linguistic features. In sung material, phoneme-monitoring and even lexical decision performance is influenced by the musical function of the sung target, even if the task does not concern the music [30, 32]. Using the musical priming paradigm, the last chord of 8-chord sequences defined the target and acted either as a strongly expected chord (the tonic) or a less expected chord (the subdominant) [30]. The four tones of each chord were sung with synthetic phonemes (e.g., /di/, /du/, /ki/, /sa:/). Participants decided whether the target was sung on a syllable containing the phoneme /i/ or /u/. Phoneme discrimination was better for strongly expected tonic targets than for subdominant targets. The finding suggests that processing of musical and phonetic information is not independent, but interacts at some stage of processing.

This interaction does not require that syllabic and musical information are combined into the same acoustic signal, but also occurs for spoken syllables that are presented in synchrony to the musical sequences in the contralateral ear [51]. It does not even require to be in the same modality: musical expectations influence visual syllable identification, when the syllables are presented in synchrony to the expected versus less-expected chords [52]. In these cross-modal experiments, participants listen to the musical sequences (ending on related or less-related chords) as background music while making speeded identification judgments on visually displayed syllables. Syllable identification was faster when the simultaneously presented chord was the related tonic chord than when it was the less-related subdominant chord.

Further experiments extended this cross-modal influence to the processing of visually displayed geometric forms [52]. This finding thus suggests that the initially reported data pattern is not specific to music and language, but attentional processes might define a common underlying process linked to expectations and temporal integration. This hypothesis is based on the dynamic attention theory proposed by Jones [53]: musical structures guide listener's attention over time and attentional resources are increased for the tonic chord (i.e., functioning as a tonal accent). These increased attentional resources would thus benefit to the simultaneous visual processing. This finding further suggests dynamic attentional resources that are shared by auditory and visual modalities.

3.3 Expectations and Musical Expressivity

Listeners' musical expectations do not only influence processing speed, but they have been attributed a role for expressiveness and emotion evoked by music [19, 54]. Based on knowledge about musical structures and relations, listeners develop expectations about future events. These expectations are not necessarily directly satisfied, but might be temporarily blocked. From this play between violations, disruptions and resolutions of expectations raise meaningful and expressive moments in music. With this role of expectations in mind, the repeated listening of a musical piece raises "Wittgenstein's puzzle" (as named in [4]): how can a well-known familiar piece be pleasant and expressive when we know exactly what will come next? Together with Jackendoff [55] and Meyer [19], Dowling and Harwood [4] propose to attribute schematic expectations a role at a subconscious level, allowing a violation of schematic expectations even if no surprise occurs at a conscious level (e.g., we remain surprised even if we know exactly that a deceptive cadence will occur).

Musical priming data provide some evidence for the automaticity of schematic expectations and their resistance to 'knowing what to come'. For single-chord contexts, Justus and Bharucha [56] opposed schematic expectations for related targets to veridical expectations for unrelated targets, which had been induced by various experimental conditions. The influence of schematic expectations on target processing was always stronger: musically related targets were processed faster even when unrelated chord pairs occurred more often or were preceded immediately by the to-be-processed target pair. Recently, we investigated with longer musical contexts and finer tonal comparisons in how far the automatic expectations based on listener's schematic knowledge can be influenced by veridical expectations and repetition priming [57]. In two-phase experiments, familiarization with a less-expected musical structure (via repeated processing) did not reverse the response time patterns. Only the exact repetition of the same sequences succeeded in decreasing, but not eliminating the processing cost of less-related targets in comparison to related targets. This behavioral data set on repeated processing of unexpected endings is in agreement with previously reported Evoked-Related Potential (ERP) data on incongruent endings (i.e., strong violations with out-of-key tones). The evoked potential linked to the expectancy violation (i.e., a late positive component peaking around 500-600ms, LPC, see 3.2.2) decreased, but persisted with repeated presentations [58].

The resistance of schematic expectations to veridical expectations provides an important element for our comprehension of musical expressiveness. Expectations for future musical events seem to be developed automatically and are not influenced by the experimental design or previously encountered exemplars. Each encountering of a musical structure seems to be like a new processing that is based on automatically developed expectations derived from tonal schematic knowledge. Lerdahl [59] describes musical forces and motion in musical space as source of musical emotion. He states "a melody or chord progression does not simply follow the inertial path of least resistance. It would be dull and would quickly come to stop unless enlivened by motion away from places that pull it towards rest" (p. 371). The priming data suggest that the tonic chord is a strongly expected event, independently of repetition or context, and that it would represent an attractive resting point. Musical sequences do not directly jump to the most expected events and thus create tension patterns, notably in relation to the tonic as the most stable, central event. The interplay between listeners' automatic schematic expectations and the realization of the musical structures would thus give rise to tension-relaxation patterns, to musical expressiveness and emotion.

4 A Connectionist Model of Nonmusician Listeners: Learning, Representation, Perception

As shown in music cognition research (see sections 2 and 3), listeners have acquired knowledge about the tonal system and its underlying regularities. Different models of mental representation have been proposed for musical knowledge, with parsimonious models based on few dimensions being preferred. Proposed models use geometric configurations ([60]; see [6] for a review), are based on either behavioral data [7, 61] or music theory [62, 63] or use artificial neural networks [11, 64]. The advantage of the artificial neural networks is not only that the model can arise from simple exposition to music, thus simulating tonal acculturation of nonmusician listeners, but also to have the possibility to present experimental musical material to the model in order to simulate the perception of nonmusician listeners.

4.1 A Hard-Wired Representation of Tonal Knowledge

Bharucha [64] proposed a connectionist account of tonal knowledge representation. In the MUSACT model (i.e., *MUSical ACTivation)*, tonal knowledge is conceived as a network of interconnected units. The units are organized in three layers corresponding to tones, chords, and keys. Each tone unit is connected to the chords of which that tone is a component. Analogously, each chord unit is connected to the keys of which it is a member.

Musical relations emerge from the activation that reverberates via connected links between tone, chord and key units. This reverberation is comparable to interactive activation processes used in word recognition models to simulate knowledge-driven influences [65, 66]. When a chord is played to MUSACT, the units representing the sounded component tones are activated and activation reverberates between the layers until equilibrium is reached (see [34, 64] for more details). The emerging activation patterns reflect tonal and harmonic hierarchies of the established key: for example, units representing harmonically related chords are activated more strongly than units representing unrelated chords. The context dependency of musical events in the tonal system is thus not stored explicitly for each of the different keys, but emerges from activation spreading through the network. The activation levels are interpreted as relative levels of expectation for future events: the more a chord unit is activated, the more the represented chord is expected and the more its processing should be facilitated. The model's architecture allows the testing of experimental material and the generation of predictions for music perception in human listeners. The model has been tested for a set of musical priming data investigating listeners' musical expectations in short and long contexts. The activation levels of the units representing the target chords in related versus unrelated contexts simulate behavioral data: activation levels were higher for related targets than for unrelated targets (see [34, 64] for details).

The MUSACT model proposes a parsimonious representation of tonal knowledge: tones and chords are presented once and not repeatedly for each tonality to reflect the contextual dependency. The change of an event's tonal function is reflected in the activation pattern, and thus emerging from the network's architecture associated with reverberation and accumulation of activation over time. However, the model is hardwired and based on music theoretic constraints. It does not simulate tonal acculturation processes to show in how far this architecture is also plausible from a learning perspective.

4.2 A Learned Representation of Tonal Knowledge

A strong advantage of artificial neural networks (e.g., connectionist models) is their capacity to adapt in such a way that representations, categorizations or associations between events can be learned. Connectionist models have the characteristic that 1) rules governing the material are not explicit, but emerge from the simultaneous satisfaction of multiple constraints represented by the connections, and 2) these constraints can be learned by repeated exposure. The MUSACT model takes advantage of the first characteristic of connectionist models. In [11], we take advantage also of the second one to simulate tonal knowledge acquisition in nonmusician listeners. For this purpose, unsupervised learning algorithms seem to be well suited: they extract statistical regularities via passive exposure and encode events that often occur together [67-70]. Self-organizing maps [68] are one version of unsupervised learning algorithms that leads to a topological organization of the learned information.

To simulate tonal acculturation, a hierarchical network composed of two self-organizing maps was exposed to short musical sequences (i.e., chord sequences). After learning, the connections in the network have changed and the units have specialized for the detection of chords and keys (the input layer coded the tones present in the input material[3]). The learned architecture is associated with a spreading activation process (as used in MUSACT) to simulate top-down influences on the activation

[3] Additional simulations integrating harmonic and subharmonic information [71] to the input pattern lead to different connection patterns, but after reverberation the activation patterns highly correlated with those of the models based on the simple input coding.

patterns. Interestingly, the learned connections and the activation patterns after reverberation mirror the outcome of the hardwired network MUSACT, which has been conceived as an idealized end-state of implicit learning processes (see [11]).

4.3 Simulating Perception of Tones, Chords and Keys

In order to be compelling, a cognitive model of music perception should not only simulate the internalization of Western pitch regularities via mere exposure, but should also simulate the behavior of listeners after having adapted to Western tonal music. The learned neural network architecture was tested for its capacity to simulate a set of empirical data on music perception. The experimental material was presented to the model[4] and the activation levels of network units were interpreted as levels of tonal stability. The more a unit (i.e., a chord unit, a tone unit) is activated, the more stable the musical event is in the corresponding context. For the experimental tasks, it was hypothesized that the level of stability affects performance (e.g., a more strongly activated, stable event is more expected or judged to be more similar to a preceding event). Overall, the simulations showed that activations in the trained self-organizing network mirror data of human participants in tonal perception experiments. The model succeeded in simulating data obtained for perceived relations between chords [16, 26, 28, 33, 72-74], between keys [14, 61] and also between tones, even if it was trained with chords only and not with melodies [7, 17, 61]. This outcome suggests the level of activations in tone, chord and key units as a single unifying concept for human performance in different perceptual tasks.

A key-finding tool. The activation levels of tone and chord units are used to simulate the perception of tones and chords in tonal contexts. The activation levels of key units can serve as a key-finding tool: without additional calculations, the key is emerging from the overall network activation. The rationale of the simulations for key perception is comparable to the simulations for tone and chord perception: the musical sequences are presented to the model and the activation levels of the key units are read out. The key unit with the maximum activation represents the key the most strongly induced by the network at that time point. When a F Major chord followed by a G Major chord is presented to the model, the most strongly activated key unit is F Major after the first chord and C Major after the second chord. As predicted by music theory, the two-chord sequence instills the C Major key. Note that the model has some independence of the stimulus encountered and inferred the key as an abstract structure (i.e., the C Major chord, the tonic, was not presented). As for this chord pair, it is possible to use the model for longer chord sequences and to track the instilled key over time (see [11] for details). The tested sequences (used by [61] for human listeners) are without modulation or contain direct and close modulation *versus* distant and remote modulations. The key-tracking over time by the model can then be compared to the judgments of human listeners. Several similarities emerge: for example, the positioning of the pivot chord in the sequence, the progressive moving through the cycle of fifths or the detection of a key without having heard the tonic yet. In sum,

[4] For event sequences, activation due to each event is accumulated and weighted according to recency [64]. The total activation of a unit is thus the sum of the stimulus activation, the phasic activation accumulated during reverberation and the decayed activation due to previous events.

the key layer of the connectionist model reveals an emerging property of key tracking. However, this key-finding tool is currently restricted to major keys. Future developments of the model thus need to include the minor keys to fully allow the exploitation of this emerging property of key-finding.

5 Implicit Learning of Tonal Knowledge and of New Musical Knowledge

Implicit learning processes enable the acquisition of highly complex information and without complete verbalizable knowledge of what has been learned [75]. Two examples of highly structured systems in our environment are language and music. Listeners become sensitive to the underlying regularities just by mere exposure to linguistic and musical materials in everyday life. The implicitly acquired knowledge then influences perception and interaction with the environment. Tonal acculturation is one example of the cognitive capacity to become sensitive to regularities in the environment. Francès [5] was one of the first underlining the importance of statistical regularities in music for tonal acculturation, suggesting that mere exposure to musical pieces is sufficient to acquire tonal knowledge, even if it remains at an implicit level. In music cognition domain, numerous research has provided evidence for nonmusicians' knowledge about the tonal system (see sections 2 and 3).

This capacity of the cognitive system is studied in the laboratory with artificial material containing statistical structures. For the auditory domain, implicit (or statistical) learning studies use structured material that is either based on artificial grammars (i.e., finite state grammars) or artificial language systems (i.e., [76-79]). The present section gives two examples with artificial structures using musical timbres and tones. Section 6 presents the application of the implicit learning approach to contemporary music, thus getting closer to ecologically valid material to which we are exposed.

5.1 Influence of Acoustic Similarities on the Learning of Statistical Regularities: Implicit Learning with Timbral Structures

Most implicit learning studies using auditory materials have focused on the statistical regularities and applied a random attribution of the sounds to the sequences. Some studies consider the acoustical characteristics of the sound, such as prosodic cues in verbal material [79-81] or acoustical similarities in non-verbal material [82]. The goal is to test whether the relation between the statistical regularities and regularities inherent to the material could influence learning: conflicting information might hinder statistical learning, while converging information might facilitate learning. Notably, tonal acculturation might represent a beneficial configuration: musical events appearing frequently together are also linked acoustically since they share (real and virtual) harmonics.

To investigate whether convergence with acoustical features represent a facilitatory or even necessary condition for statistical learning, Tillmann and McAdams [82] systematically manipulated acoustical similarities between musical timbres so that they either underline the statistical regularities of the timbre sequences, contradict these regularities or are neutral to them. The statistical regularities were defined as in artificial language studies (see [79]). Based on a restricted set of elements (syllables or here musical timbres), groups of three elements define units (three syllables define artificial words, three timbres define timbre-triplets). These units are chained together in a continuous sequence without silences in between. The transition probabilities between elements inside a unit are high, while transition probabilities between elements crossing units are low (i.e., a unit can be followed by one of the six other units)[5]. If listeners become sensitive to these statistical regularities, they should be able to extract the triplets of timbres from the continuous sequence, just as listeners are able to extract words from a syllable-stream of an artificial language (e.g., [79]). In [82], the sequences were constructed in such a way that the acoustical dissimilarities between timbres potentially created perceptual segmentations that either supported (S1) or contradicted (S2) the statistical regularities or were neutral with respect to them (S3).

To manipulate the acoustical similarities/dissimilarities, musical timbres were selected from the timbre space defined by [83]. Timbre is a multidimensional set of auditory attributes that is based on temporal and spectral features of sounds (cf. also [84-86]). Based on perceived dissimilarity judgments, a multidimensional analysis revealed a three-dimensional spatial structure in which the synthesized timbres were placed and the distances between timbres reflect the perceived dissimilarities among them. For example, the horn timbre is close in space to the trombone timbre (both brass instruments), but is distant from that of the vibraphone (a percussion instrument). For S1, the timbres of the triplets were chosen in such a way that the distances between adjacent timbres inside the triplets were small, but the distances between the last timbre of any given triplet and the first timbre of all other triplets (across boundaries) in the sequence were large. For S2, the distances between timbres inside the triplets were large, but the distances between timbres of two successive triplets (across boundaries) were small. For S3, mean distances between timbres inside the triplets were equal to mean distances between timbres of two successive triplets. In S1, the triplets were thus defined by statistical cues and by abrupt acoustical changes between triplets. In S2 and S3, the triplets were only defined by statistical cues, while in S2 the acoustical similarities were out of phase with the statistical boundaries. For the three sequences, the transition probabilities inside the triplets and across triplet boundaries were identical, and the same set of timbres was used.

The experiments consisted of two phases: an exposition phase and a testing phase. In the learning group, participants first listened to the continuous timbre sequence without being told about the triplets. In the testing phase, participants had then to distinguish statistical units from new units. This test performance was then compared to a control group that was lacking the exposition phase and was directly working on the test phase. The comparison of test performance between learning and control groups allows estimating the amount of learning as well as the initial biases influencing the responses without prior exposition.

The data of the learning group in comparison to the control group revealed no interaction between sequence type (S1, S2, S3) and amount of learning: performance increased by the same amount for the three sequences. After exposure, participants

[5] The transition probability of B given A is calculated as the frequency of the pair AB divided by the absolute frequency of A [79].

were better in chosen the existing triplets over other associations of three timbres. Additionally, performance reflected an overall preference for acoustically similar timbre triplets (in S1) over dissimilar timbre triplets (S2). This outcome extends previous data from the domain of implicit learning to complex nonverbal auditory material. It further suggests that listeners become sensitive to statistical regularities despite similarities or differences among the acoustical surface characteristics in the material. The surface characteristics only affected grouping and overall preference bias for the different materials.

This data set suggests that tonal acculturation does not necessarily need the convergence between statistical and acoustical regularities. Supporting evidence can be found in acculturation to Arabic music, which is lacking the convergence between statistical and acoustic features [87]. Together with the implicit learning study on twelve-tone music ([88]; see section 6), the data emits the rather encouraging hypothesis about the possibility to learn regularities of new musical styles independently of acoustical features.

5.2 Implicit Learning of Regularities of an Artificial Tone System: Grammaticality Judgments and Tone Expectations

The seminal studies by Reber [77] used artificial grammars to study implicit learning processes. A finite state grammar and a restricted set of letters were used to create letter strings. In a typical experimental setting, participants are first exposed to stimuli that are based on a finite-state grammar, without being told about the grammatical structure. After exposure, participants are informed about the stimuli's grammaticality and are required to classify novel sequences as grammatical or ungrammatical. Performance is generally above chance-level, without (or only little) verbalizable knowledge of the regularities underlying the letter sequences.

For the adaptation of this paradigm in the auditory domain, the letters of the artificial grammars are replaced by sine waves [76], musical timbres (e.g., gong, trumpet, piano, violin, voice in [89]) or environmental sounds (e.g., drill, clap, steam in [90, 91]). The basic experimental design remained the same as in the original studies using letters. In Altmann et al. [76], for example, letters were translated into tones (i.e., generated with sine waves) by using a random mapping of tone frequencies to letters (e.g., the letter M became the tone C), and participants' performance was as high when trained and tested with letter strings as with tone sequences. These studies provided evidence that implicit learning processes also operate on auditory sequences and that the simple exposure to sequences generated by a statistical system allows participants to distinguish sequences that break the rules.

In two recent studies, we used an artificial grammar with tones (i.e., creating tone sequences) and tested implicit learning of these regularities with a) explicit, direct grammaticality judgments and b) an adaptation of the priming paradigm to investigate whether musical expectations can be developed with the newly acquired knowledge. The novelty of our approach was to test listeners with new grammatical items that were opposed to ungrammatical items containing very subtle violations. Notably, the tones creating the ungrammaticality in the sequence belonged to the tone set of the grammar and they respected frequency distributions of tones, bigrams, and melodic contour (i.e., as defined for the grammatical sequences). For these tone structures, participants' grammaticality judgments were above chance-level after an exposure phase. This outcome suggests that listeners became sensitive to the regularities underlying the used artificial grammar of tones [92].

In a second study, we combined implicit learning and priming paradigms to investigate whether newly acquired structure knowledge allows listeners to develop perceptual expectations for future events [93]. Participants were first exposed to structured tone sequences (based on the artificial grammar), and made then speeded judgments on a perceptual feature of target tones in new sequences. The priming task was adapted from musical priming research (see section 3 and [94]) and required participants to judge whether the target tone was played either in-tune or out-of-tune. Most importantly, the target tone either respected the artificial grammar structure or violated it by creating subtle ungrammaticalities (as in [92]). In this priming task, grammatical tones were processed faster and more accurately than ungrammatical ones. This processing advantage was not observed for a control group, which was lacking the exposure phase to the grammatical tone sequences. This finding suggests that the acquisition of new structure knowledge allows listeners to develop auditory expectations that influence single event processing. It further promotes the priming paradigm as an implicit access to acquired artificial structure knowledge studied in the lab. A recent extension of this experimental approach showed similar cross-modal influences as had been observed for the musical material (see section 3.3): response times for visual syllable identification were faster when a grammatical tone was played at the same time than when a tone was played that created an ungrammaticality.

These studies with artificial tone structures imitate the phenomenon of tonal acculturation inside the lab: Nonmusicians acquire implicit knowledge of the Western tonal system by mere exposure to musical pieces obeying this system. The beneficial influence of auditory expectations on the processing of expected events can arise after short exposure to a structured tone system in the laboratory. Based on this finding, we make the hypothesis that musical expectations and their influence on processing efficiency can also occur after exposure to new musical systems (e.g., the twelve-tone music tested by [88], see section 6). Furthermore, the experiment using timbres presented above [82] further suggests that these processes might occur independently of acoustical surface characteristics and their combination to statistical regularities of the musical system.

6 Learning, Perception and Expectations in Other Musical Systems

Listeners who are familiarized to the music of their culture do not perceive a disorganized superposition of sounds or groups of sounds, but they perceive coherent melodic lines, they develop expectations and anticipate possible continuations and endings of a musical piece. Research in music cognition analyzes how listeners succeed in these processes, and aims to specify listeners' knowledge about the musical system, its acquisition, structure and influence in perception and performance.

Most research on music perception and cognition has focused on the Western tonal musical system of the 18^{th} and 19^{th} centuries. And even if the principal regularities of this system are used in a variety of musical styles (classical music, pop, folk, jazz etc.), this represents a restriction that needs to be redressed. Notably, a more general theory of music perception and cognition requires studying the hypothesis about learning, perception, knowledge and expectations also for other musical systems and listeners [3, 95]. Regularities between musical events also exist in other musical systems (e.g., Indian or Arabic music) and cultural learning and familiarity to these systems lead to auditory experiences different from those of naive listeners.

This section presents some research studying cognition and perception of musical systems from other cultures and of new musical systems. The overall results point out that the acculturation processes also apply to other musical systems: listeners acquire knowledge about their musical system (or to a new musical system) by mere exposure and this knowledge influences perception of musical structures. Even if these findings suggest the generality of the processes underlying learning and perception of music (and are rather encouraging for creators of new musical systems), research on musical systems other than the tonal system remains rather rare up to now. For contemporary music, one example of a perceptual investigation has been realized on a piece by Roger Reynolds (The Angel of Death). The perception of its musical structures has been investigated by a series of behavioral experiments, contrasting also the composer's intent to the listeners' understanding (see Special Issue of Music Perception, 2004, Vol. 22 (2)).

Perception of pitch regularities in other musical systems. The probe-tone paradigm has been used to investigate perception of scale structures in Balinese music [96] and Indian music [97]. As in the original studies on Western tonal music [61], a context was followed by one of the possible tones and listeners rated how well this tone fits into the preceding context. Both studies compared the perception of the scale structures by naïve listeners and by acculturated, native listeners. For Indian music, for example, the data patterns of both groups of listeners showed sensitivity to the sensory information present in the context. However, only the Indian listeners (but not the North American listeners) showed the perception of fine-graded musical features that were independent of the tones presented in the context [97]. This outcome can be interpreted in terms of musical knowledge that Indian listeners have acquired by mere exposure, while American listeners were missing this acculturation process.

Converging evidence has been reported with segmentation tasks for Arabic music: both Arabic and European listeners use salient surface features for segmentation (i.e., pauses, register changes), but only Arabic listeners use cues based on subtle modal changes [87]. Finally, differences between novice and expert listeners have also been reported for Finnish spiritual folk hymns and North Sami yoiks [98, 99].

Learning and perception of time structures. Numerous research conducted with Western listeners (i.e., Western European and North American listeners) have shown that perception and production of weakly metric rhythms is less accurate than perception and production of strongly metric rhythms (e.g., [100, 101]). Simple integer ratios in general are easier to perceive than complex ratios [102], leading to the hypothesis of more complex cognitive processes necessary for the processing of complex ratio (e.g., 2:3) *versus* simple ratio meters (e.g., 1:2).

However, Hannon and Trehub [103] recently showed the importance of acculturation in the perception of metrical patterns. Meters with simple ratio predominate in Western music, while meters with complex ratio are common in other musical cultures, as for example Macedonian music. While North American adults showed weaker performance for the complex-meter patterns than the simple-meter patterns, Macedonian and Bulgarian adults performed equally well with both patterns. The hypothesis of the importance of exposure (versus the complexity of processing or cognitive predisposition for simple meter processing) received further support by additional infant experiments: 6-month-old North American infants performed equally well for both metric patterns, thus being able to process even complex patterns. However, by the age of 12-months, North American infants performed like North American adults with a bias for simple patterns [103, 104]. In contrast to adults, this bias of infants was reversible with simple training by exposure to complex patterns, thus suggesting some sensitive period for the acquisition of temporal structures. In sum, this example on temporal perception illustrates the pitfalls of Western-focused research and the importance of testing the perception of other musical systems for both naïve and native listeners.

Simulating learning and perception of other musical systems. Artificial neural networks have been used for the simulation of the learning and perception not only of Western tonal music, as exposed above, but also of other musical systems. Krumhansl et al. [98, 99] used self-organizing maps to simulate melodic expectancies by experts of North Sami yoiks versus experts of Finnish folk songs or Lutheran hymns. Different models were trained on the different systems and their predictions for yoiks were compared with behavioral data obtained for human listeners with various expertise.

Two further examples of the use of connectionist models to simulate learning of knowledge and its influence on music perception have been proposed by Bharucha and Olney [105]. A connectionist model (i.e., an auto-associative net) is trained with Indian *rags* (those used in the study by [97]). After learning, the network fills in missing tones of the scale and the authors make the link to faster processing of expected events (i.e., tones of the activated scale pattern). In addition, a trained connectionist model can serve to simulate the perceptual filter or bias created by the knowledge of one musical system on the perception of another system: A connectionist model is first exposed to the regularities of Western tonal music. Once learning has occurred, an Indian *rag* is presented to the model. The simulations show that the model assimilates the *rag* to the learned major and minor keys. The model thus shows the assimilation processes of Indian structures to the Western schemata learnt previously.

Implicit learning of regularities in new musical systems. Tonal acculturation is an example of implicit learning processes on material encountered in everyday life and leading to nonmusicians' implicit knowledge about the Western tonal musical system. The few studies on the perception of music of other cultures by native listeners suggest similar acculturation processes for exposure to music of other cultures (on both pitch and time dimensions). Implicit learning research studies the strengths and limits of this cognitive capacity in the laboratory with artificial material. Bridging the gap between complexity of real life learning and artificiality of experimental material, Bigand and collaborators [88] investigated the implicit learning of twelve-tone music in the laboratory. This atonal musical system is based on a tone row, the ordered arrangement of the 12 tones of the chromatic scale (forming a basic rule of 12-tone musical grammar initially proposed by Schoenberg [106], 1941). One piece of music is based on one row and its possible transformation. Historically, the proposition of this tone system that broke with the concepts of tonal structures and hierarchy has led to considerable debate about whether listeners can understand these new structures. The researchers have brought this question into the lab with the implicit learning paradigm. First, listeners were exposed to musical pieces composed with a specific 12-tone row. In the test phase, participants listened to new excerpts presented by pair and had to select the excerpt that "was composed by the same composer". More specifically, one excerpt was based on the same row as in the exposition phase, and the other excerpt on a different row. Participants (musicians and nonmusicians) performed above chance in this test, even if they were very uncertain of their responses. Moreover, a control group, which had been exposed to excerpts based on both rows, did not differ from chance. This experiment suggests that the listeners became sensitive to the specific atonal structures in the exposure phase despite the complexity of the material [88].

References

1. Palmer, C.: Sequence memory in music performance. Current Directions in Psychological Science 14, 247–250 (2005)
2. Repp, B.H.: Sensorimotor synchronization: A review of the tapping literature. Psychonomic Bulletin & Review 12, 969–992 (2005)
3. Stevens, C., Byron, T.: Universals in music processing. In: Hallmam, C.T. (ed.) Oxford Handbook of Music Psychology. Oxford (2008)
4. Dowling, W.J., Harwood, D.L.: Music Cognition. Academic Press, Orlando (1986)
5. Francès, R.: La perception de la musique, 2nd edn. Vrin, Paris (1958)
6. Krumhansl, C.L.: Cognitive foundations of musical pitch. Oxford University Press, New York (1990)
7. Krumhansl, C.L.: The psychological representation of musical pitch in a tonal context. Cognitive Psychology 11(3), 346–374 (1979)
8. Bharucha, J.J.: Anchoring effects in music: The resolution of dissonance. Cognitive Psychology 16(4), 485–518 (1984)
9. Budge, H.: A study of chord frequencies. Teacher College (1943)
10. Bigand, E., Poulin-Charronnat, B.: Are we all experienced listeners? Cognition 100, 100–130 (2006)
11. Tillmann, B., Bharucha, J.J., Bigand, E.: Implicit learning of tonality: a self-organizing approach. Psychol Rev. 107(4), 885–913 (2000)
12. Hébert, S., Peretz, I., Gagnon, L.: Perceiving the tonal ending of tune excerpts: The roles of pre-existing representation and musical expertise. Canadian Journal of Experimental Psychology 49, 193–209 (1995)
13. Bartlett, J.C., Dowling, W.J.: Recognition of transposed melodies: a key-distance effect in developmental perspective. J. Exp. Psychol. Hum. Percept. Perform 6(3), 501–515 (1980)
14. Cuddy, L.L., Thompson, W.F.: Perceived key movement in four-voice harmony and single voices. Music Perception 9, 427–438 (1992)
15. Bigand, E.: Perceiving musical stability: the effect of tonal structure, rhythm, and musical expertise. J. Exp. Psychol. Hum. Percept. Perform 23(3), 808–822 (1997)
16. Bharucha, J.J., Krumhansl, C.L.: The representation of harmonic structure in music: hierarchies of stability as a function of context. Cognition 13(1), 63–102 (1983)

17. Dowling, W.J.: Scale and contour: Two components of a theory of memory for melodies. Psychological Review 85(4), 341–354 (1978)
18. Tillmann, B., Bigand, E., Madurell, F.: Local versus global processing of harmonic cadences in the solution of musical puzzles. Psychological Research/Psychologische Forschung 61(3), 157–174 (1998)
19. Meyer, L.B.: Emotion and Meaning in Music. University of Chicago Press, Chicago (1956)
20. Carlsen, C.: Musical expectancy: Some perspectives. Council for Research in Music Education 71, 4–14 (1982)
21. Carlsen, C.: Some factors which influence melodic expectancy. Psychomusicology 1(1), 12–29 (1981)
22. Schmuckler, M.A.: The performance of global expectations. Psychomusicology 9, 122–147 (1990)
23. Schmuckler, M.A.: Expectation in music: Investigation of melodic and harmonic processes. Music Perception 7, 109–150 (1989)
24. Schmuckler, M.A., Boltz, M.G.: Harmonic and rhythmic influences on musical expectancy. Percept Psychophys 56(3), 313–325 (1994)
25. Neely, J.H.: Semantic priming effects in visual word recognition: A selective review of current findings and theories. In: Besner, D., Humphreys, G.W. (eds.) Basic processes in reading: Visual word recognition, pp. 264–336. Lawrence Erlbaum, Mahwah (1991)
26. Tillmann, B.: Implicit investigations of tonal knowledge in nonmusician listeners. Annals of the New York Academy of Sciences 1060, 100–110 (2005)
27. Bharucha, J.J., Stoeckig, K.: Reaction time and musical expectancy: priming of chords. J. Exp. Psychol. Hum. Percept. Perform 12(4), 403–410 (1986)
28. Bigand, E., Pineau, M.: Global context effects on musical expectancy. Percept Psychophys 59(7), 1098–1107 (1997)
29. Tillmann, B., Bharucha, J.J.: Effect of harmonic relatedness on the detection of temporal asynchronies. Perception & Psychophysics 64(4), 640–649 (2002)
30. Bigand, E., et al.: The effect of harmonic context on phoneme monitoring in vocal music. Cognition 8(1), B11–B20 (2001)
31. Tillmann, B., et al.: Influence of harmonic context on musical timbre processing. European Journal of Cognitive Psychology 18, 343–358 (2005)
32. Poulin-Charronnat, B., et al.: Musical structure modulates semantic priming in vocal music. Cognition 94, B67–B78 (2005)
33. Tillmann, B., Bigand, E., Pineau, M.: Effects of global and local contexts on harmonic expectancy. Music Perception 16(1), 99–117 (1998)
34. Bigand, E., et al.: Effect of global structure and temporal organization on chord processing. Journal of Experimental Psychology: Human Perception and Performance 25(1), 184–197 (1999)
35. Bigand, E., et al.: Cognitive versus sensory components in harmonic priming effects. Journal of Experimental Psychology: Human Perception and Performance 29(1), 159–171 (2003)
36. Tillmann, B., et al.: Tonal centers and expectancy: facilitation or inhibition of chords at the top of the harmonic hierarchy? Journal of Experimental Psychology: Human Perception & Performance (in press)
37. Marmel, F., Tillmann, B., Dowling, W.J.: Tonal expectations influence pitch perception (manuscript submitted for publication, 2007)

38. McAdams, S.: Contraintes psychologiques sur les dimensions porteuses de formes en musique. In: McAdams, S., Deliege, I. (eds.) La musique et les sciences cognitives, pp. 257–284. Bruxelles, Mardaga (1989)
39. Boltz, M.G.: The generation of temporal and melodic expectancies during musical listening. Perception & Psychophysics 53, 585–600 (1993)
40. Tillmann, B., Lebrun-Guillaud, G.: Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. Psychological Research 70, 345–358 (2006)
41. Peretz, I., Kolinsky, R.: Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. Quarterly Journal of Experimental Psychology 46A, 301–327 (1993)
42. Boltz, M.G.: Some structural determinants of melody recall. Mem Cognit 19(3), 239–251 (1991)
43. Boltz, M.G.: Perceiving the end: Effects of tonal relationships on melodic completion. Journal of Experimental Psychology: Human Perception and Performance 15, 749–761 (1989)
44. Palmer, C., Krumhansl, C.L.: Independent temporal and pitch structures in determination of musical phrases. J. Exp. Psychol. Hum. Percept. Perform 13(1), 116–126 (1987)
45. Bigand, E.: The influence of implicit harmony, rhythm and musical training on the abstraction of "tension-relaxation schemes" in a tonal musical phrase. Contemporary Music Review 9, 128–139 (1993)
46. Jones, M.R., Boltz, M.: Dynamic attending and responses to time. Psychological Review 96, 459–491 (1989)
47. Dowling, W.J., Lung, K.M., Herrbold, S.: Aiming attention in pitch and time in the perception of interleaved melodies. Percept Psychophys 41(6), 642–656 (1987)
48. Lerdahl, F., Jackendoff, R.: A generative Theory of Tonal Music, vol. 368. The MIT press, Cambridge (1983)
49. Peretz, I., Morais, J.: La musique et la modularité. In: McAdams, S., Deliege, I. (eds.) La musique et les sciences cognitives, pp. 393–414. Bruxelles, P. Mardaga (1989)
50. Pfordresher, P.Q.: The role of melodic and rhythmic accents in musical structure. Music Perception 20(4), 431–464 (2003)
51. Hoch, L., Tillmann, B.: Effect of tonal relatedness on spoken syllable identification to the contralateral ear (manuscript in preparation)
52. Escoffier, N., Tillmann, B.: The tonal function of a task-irrelevant chord modulates speed of visual processing. Cognition (in press)
53. Jones, M.R.: Dynamic pattern structures in music: Recent theory and research. Perception and Psychophysics 41, 621–634 (1987)
54. Meyer, L.B.: On rehearing music. In: Meyer, L.B. (ed.) Music, the arts and ideas, pp. 42–53. Chicago University Press, Chicago (1967)
55. Jackendoff, R.: Musical parsing and musical affect. Music Perception 9, 199–230 (1991)
56. Justus, T.C., Bharucha, J.J.: Modularity in musical processing: the automaticity of harmonic priming. J. Exp. Psychol. Hum. Percept. Perform 27(4), 1000–1011 (2001)
57. Tillmann, B., Bigand, E.: Musical priming: Schematic expectations resist repetition priming. In: 8th International Conference of Music Perception and Cognition. Evanston, Chicago (2004)
58. Faita, F., Besson, M.: Electrophysiological index of musical expectancy: Is there a repetition effect on the event-related potentials associated with musical incongruities? In: Third International Conference for Music Perception and Cognition, ESCOM, Liege (1994)

59. Lerdahl, F.: Two ways which music relates the world. Music theory spectrum 25, 367–373 (2003)
60. Shepard, R.N.: Geometrical approximations to the structure of musical pitch. Psychol. Rev. 89(4), 305–333 (1982)
61. Krumhansl, C.L., Kessler, E.J.: Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. Psychol. Rev. 89(4), 334–368 (1982)
62. Lerdahl, F.: Tonal pitch space. Music Perception 5(3), 315–349 (1988)
63. Lerdahl, F.: Pitch-space journeys in two Chopin Preludes. In: Jones, M.R., Holleran, S. (eds.) Cognitive bases of musical communication, APA, pp. 171–191 (1991)
64. Bharucha, J.J.: Music cognition and perceptual facilitation: A connectionist framework. Music Perception 5(1), 1–30 (1987)
65. McClelland, J.L., Rumelhart, D.E.: An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. Psychological Review 86, 287–330 (1981)
66. Seidenberg, M.S., McClelland, J.L.: A distributed, developmental model of word recognition and naming. Psychological Review 96, 523–568 (1989)
67. Grossberg, S.: Some networks that can learn, remember and reproduce any number of complicated space-time patterns. Studies in Applied Mathematics 49, 135–166 (1970)
68. Kohonen, T.: Self-Organizing Maps. Springer, Heidelberg (1995)
69. Rumelhart, D.E., Zipser, D.: Feature discovery by competitive learning. Cognitive Science 9, 75–112 (1985)
70. von der Malsberg, C.: Self-organizing of orientation sensitive cells in the striate cortex. Kybernetic 14, 85–100 (1973)
71. Parncutt, R.: Harmony: A psychoacoustical approach. Springer, Heidelberg (1989)
72. Krumhansl, C.L., Bharucha, J.J., Kessler, E.J.: Perceived harmonic structures of chords in three related keys. Journal of Experimental Psychology: Human Perception and Performance 8, 24–36 (1982)
73. Patel, A.D., et al.: Processing syntactic relations in language and music: an event-related potential study. J. Cogn. Neurosci. 10(6), 717–733 (1998)
74. Tekman, H.G., Bharucha, J.J.: Implicit knowledge versus psychoacoustic similarity in priming of chords. Journal of Experimental Psychology: Human Perception and Performance 24(1), 252–260 (1998)
75. Seger, C.A.: Implicit learning. Psychological Bulletin 115, 163–169 (1994)
76. Altmann, G.T.M., Dienes, Z., Goode, A.: Modality independence of implicitly learned grammatical knowledge. Journal of Experimental Psychology: Learning, Memory, and Cognition 21(4), 899–912 (1995)
77. Reber, A.S.: Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior 6, 855–863 (1967)
78. Reber, A.S.: Implicit learning and tacit knowledge. Journal of Experimental Psychology: General 118, 219–235 (1989)
79. Saffran, J.R., Newport, E.L., Aslin, R.N.: Word segmentation: The role of distributional cues. Journal of Memory and Language 35(4), 606–621 (1996)
80. Thiessen, E.D., Saffran, J.R.: When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. Developmental Psychology 39(4), 706–716 (2003)
81. Johnson, E.K., Jusczyk, P.W.: Word segmentation by 8-month-olds: When speech cues count more than statistics. Journal of Memory and Language 44(4), 548–567 (2001)

82. Tillmann, B., McAdams, S.: Implicit Learning of musical timbre sequences: statistical regularities confronted with acoustical (dis)similarities. Journal of Experimental Psychology: Learning, Memory & Cognition 30, 1131–1142 (2004)
83. McAdams, S., et al.: Perceptual scaling of synthesized musical timbres: Common dimensions, specificities and latent subject classes. Psychological Research 58, 177–192 (1995)
84. Grey, J.M.: Multidimensional perceptual scaling of musical timbres. Journal of the Acoustical Society of America 61, 1270–1277 (1977)
85. Krumhansl, C.L.: Why is musical timbre so hard to understand? In: Nielzen, S., Olsson, O. (eds.) Structure and perception of electroacoustic sound and music, pp. 43–54. Excerpta medica, Amsterdam (1989)
86. Samson, S., Zatorre, R.J., Ramsay, J.O.: Multidimensional scaling of synthetic musical timbre: perception of spectral and temporal characteristics. Canadian Journal of Experimental Psychology 51, 307–315 (1997)
87. Ayari, M., McAdams, S.: Aural analysis of Arabic improvised instrumental music (tagsim). Music Perception 21, 159–216 (2003)
88. Bigand, E., D'Adamo, D.A., Poulin, B.: The implicit learning of twelve-tone music. In: ESCOP 2003, Granada, Spain (2003)
89. Bigand, E., Perruchet, P., Boyer, M.: Implicit learning of an artificial grammar of musical timbres. Cahiers de Psychologie Cognitive/Current Psychology of Cognition 17(3), 577–600 (1998)
90. Howard, J.H.J., Ballas, J.A.: Acquisition of acoustic pattern categories by exemplar observation. Organization, Behavior and Human Performance 30, 157–173 (1982)
91. Howard, J.H.J., Ballas, J.A.: Syntactic and semantic factors in the classification of nonspeech transient patterns. Perception & Psychophysics 28(5), 431–439 (1980)
92. Poulin-Charronnat, B., Tillmann, B., Perruchet, P.: Implicit learning of artificial grammar of tones: direct and indirect judgments (manuscript in preparation)
93. Tillmann, B., Poulin-Charronnat, B.: Auditory expectations for newly acquired material: Combining implicit learning and priming paradigms (manuscript in preparation)
94. Tillmann, B., Marmel, F.: Testing musical expectations at various positions inside a chord sequence: An adaptation of the musical priming paradigm (manuscript submitted for publication, 2007)
95. Stevens, C.: Cross-cultural studies of musical pitch and time. Acoustical Science and Technology 25, 433–438 (2004)
96. Kessler, E.J., Hansen, C., Shepard, R.N.: Tonal schemata in the perception of music in Bali and in the West. Music Perception 2, 131–165 (1994)
97. Castellano, M.A., Bharucha, J.J., Krumhansl, C.L.: Tonal hierarchies in the music of North India. Journal of Experimental Psychology: General 113, 394–412 (1984)
98. Krumhansl, C.L., et al.: Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks. Cognition 76(1), 13–58 (2000)
99. Krumhansl, C.L., et al.: Melodic expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches. Music Perception 17, 151–195 (1999)
100. Povel, D.-J., Essens, P.J.: Perception of temporal patterns. Music Perception 2, 411–440 (1985)
101. Keller, P.E., Burnham, D.K.: Musical meter in attention to multipart rhythm. Music Perception 22, 629–661 (2005)
102. London, J.: Hearing in time. Oxford University Press, New York (2004)
103. Hannon, E.E., Trehub, S.E.: Metrical categories in infancy and adulthood. Psychological Science 16, 48–55 (2005)

104. Hannon, E.E., Trehub, S.E.: Tuning in to musical rhythms: Infants learn more readily than adults. Proceedings of the National Academy of Sciences of the United States of America 102, 12639–12643 (2005)
105. Bharucha, J.J., Olney, K.L.: Tonal cognition, artificial intelligence and neural nets. Contemporary Music Review 4, 341–356 (1989)
106. Schoenberg, A.: Style and idea: Selected writings of Arnold Schonberg. Farber and Farber, London (1941)