26. Musical Syntax II: Empirical Perspectives

Marcus Pearce, Martin Rohrmeier

Efforts to develop a formal characterization of musical structure are often framed in syntactic terms, sometimes but not always with direct inspiration from research on language. In Chap. 25, we present syntactic approaches to characterizing musical structure and survey a range of theoretical issues involved in developing formal syntactic theories of sequential structure in music. Such theories are often computational in nature, lending themselves to implementation and our first goal here is to review empirical research on computational modeling of musical structure from a syntactic point of view. We ask about the motivations for implementing a model and assess the range of approaches that have been taken to date. It is important to note that while a computational model may be capable of deriving an optimal structural description of a piece of music, human cognitive processing may not achieve this optimal performance, or may even process syntax in a different way. Therefore we emphasize the difference between developing an optimal model of syntactic processing and developing a model that simulates human syntactic processing. Furthermore, we argue that, while optimal models (e.g., optimal compression or prediction) can be useful as a benchmark or yardstick for assessing human performance, if we wish to understand human cognition then simulating human performance (including aspects that are nonoptimal or even erroneous) should be the priority. Following this principle, we survey research on processing of musical syntax from the perspective of computational modeling, experimental psychology and cognitive neuroscience. There exists a large number of computational models of musical syntax, but we limit ourselves to those that are explicitly cognitively motivated, assessing them in the context of theoretical, psychological and neuroscientific research.

26.1 Computational Research

26.1.1 Foundations

Different approaches to building computer models of musical structure can be characterized, and distinguished, in terms of how expressive they are in terms of the degree of structural complexity they are capable of representing. Therefore, there is a direct link between the theoretical characterization of musical syntax (discussed in Chap. 25) and the implementation and testing of these theories as computational models of cognition, discussed here. Implementing a theory of musical syntax and processing has several potential advantages,

R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, https://doi.org/10.1007/978-3-662-55004-5_26

following well-known examples in cognitive science:

1. Implementing a theory as a computer program (which must run, generating output from the data and parameters supplied as input) ensures that it takes as little as possible for granted and any assumptions are explicitly stated [26.1–3].
2. Experiments can be run to evaluate the implemented theory by comparison of its behavior with the output expected from theoretical accounts or directly with human behavior given the same inputs [26.3, 4].

It is important to distinguish between models whose knowledge is provided to them by a human expert (in the style of good, old-fashioned artificial intelligence, GOFAI) and those that acquire knowledge about musical structure by learning (either supervised or unsupervised) from experience of music, given some predefined structural representation, the parameters of which are learned (in the tradition of machine learning), be it a neural network, a Markov model or grammatical inference. The appropriate approach depends on the goal of the modeling. However, in many cases the two approaches are complementary in that successfully learning complex representations (e.g., context-free rules) is extremely challenging but the alternative approach can result in models that do not generalize far beyond the musical domain for which they were developed (e.g., *Steedman*'s [26.5, 6] grammars for blues harmony). In the past, tasks that required context-free representations have usually been hand coded while tasks that require simpler representational relationships have been able to benefit from the flexibility of machine learning. However, methods have now been developed in computational linguistics to learn certain kinds of context-free representation [26.7]. Note that if we are interested in cognitive representations of music, there is the additional issue of the extent to which the representations in question are actually learned or inherited (i. e., innately specified) by human beings ([26.8] as, e.g., discussed in the poverty of the stimulus debate [26.9]).

As mentioned above, we must also emphasize the distinction between finding an optimal structural description of a piece of music and modeling the cognitive representation of that piece in the mind of a listener. The former bounds the latter but listeners are likely to be subject to constraints of perception and cognition (e.g., limitations of working memory load), which would prevent them reaching an optimal structural description. Note also that it is problematic to assume the existence of an *average listener* without understanding all the factors (e.g., musical training, environmental context, degree of attention etc.) that could influence the structural descriptions that listeners form. Nonetheless, it is often useful to identify theoretical bounds on structural complexity using an optimal model. Furthermore, theoretical models of musical structure can help us understand the hypothesis space that human learners are faced with when they acquire the syntactic structure of a musical style. In artificial intelligence research, we distinguish between the representation defining the search space and algorithms for traversing the space. Similarly in machine learning, we distinguish between the hypothesis space and learning mechanisms for traversing the space. In the case of musical syntax, the hypotheses correspond to potential stylistic grammars generating structural descriptions of music and we can think of the learning as traversing the space of possible grammars, specifying the parameters distinguishing those grammars along the way.

The following sections illustrate these points, using different kinds of computational models that have been proposed for understanding musical syntax.

26.1.2 Early Approaches: Pattern Processing

We begin with a review of two early approaches that are of historical importance because they laid the foundations for symbolic models that subsequently became influential. One early approach was based on the assumption that listeners use pattern induction processes to develop predictions for successive events in melodies [26.10, 11]. These models attempt to define formal languages for describing the patterns perceived by humans in temporal sequences (such as music) and use them to explain how these patterns are applied for prediction. *Simon* and *Sumner* [26.11], for example, begin with ordered alphabets for representing the range of possible values for a particular musical dimension (e.g., note names, note durations). Simon and Sumner restrict their attention to the dimensions of melody, harmony, rhythm and form and use alphabets for diatonic notes, triads, duration, stress and formal structure. The operations they consider are *same* (when the subsequent symbol is identical to the previous one) and *NEXT* (the subsequent symbol is obtained by taking the next symbol in the specified alphabet a specified number of times). Sequences of symbols may then be described more compactly as a sequence of these operations.

Deutsch and *Feroe* [26.10] extended the model of Simon and Sumner in several ways, in particular defining structures as sequences of elementary operators and sequence operators such as prime, retrograde, inversion and alternation. They apply their pattern language to various alphabets, corresponding to collections of pitches, such as the major and (natural, harmonic and melodic) minor scales, major, minor and diminished triads, and seventh chords. They argue that the pattern language facilitates processing in four ways:

1. Reduced redundancy of representation
2. Distinct alphabets may be invoked at different levels
3. Embedded sequence structures and their associated alphabets may be encoded as chunks
4. The chunking of structures allows for the representation of configurations that satisfy proximity and the differentiation of different members of the alphabet in terms of frequency.

Deutsch and Feroe propose that multiple representations may be formed by listeners who, according to the model, will tend to choose the most parsimonious. The acquisition of a representation is an ongoing process of generation and testing of multiple hypothesized structural representations [26.12, 13].

26.1.3 Markov Modeling

Another early approach to modeling musical structure was based on statistical learning and information theory, in particular using *information content* and *entropy* to measure the complexity of a musical style. It is interesting to note that information theory was applied to music as early as 1955 [26.14–17] just a few years after *Shannon*'s foundational work was published [26.18]. Typically, this approach involves representing a musical work as a sequence of symbols drawn from an alphabet (e.g., a melody might be represented as a sequence of pitch symbols, harmonic movement as a sequence of chord symbols). The learning task is to estimate a conditional probability governing the next symbol in the musical sequence, given the preceding symbols. Such models (in various guises) have been highly influential in terms of understanding predictive processing in human cognitive processing of music [26.19, 20].

It is possible to vary the length of the context used in estimating the probability, known as the *order* of the model (a zeroth-order model has no context, a first-order model a context of one symbol and so on). Usually the probabilities (i. e., the parameters of the model) are estimated through statistical analysis of a corpus of musical works. These models are also known as Markov models because the probability of an event is only dependent on the immediately previous context (the Markov property), or, in other words, the model does not take into account any nonlocal dependencies. Once a probability distribution has been generated in a particular context, the *entropy* of that distribution reflects the model's uncertainty about the following musical event before it arrives while the *information content* (the negative log probability) reflects how unexpected the next note is, once it has arrived – given a local model [26.21]. Note that the entropy and information content of a musical event or sequence are not properties of the music per se, but properties of the music from the perspective of an underlying model.

It is interesting that one of the first nonmilitary applications of early computers was software to generate music incorporating grammatical representations of musical styles corresponding to probabilistic Markov models [26.22]. Early work using these models tended to focus on fixed, low-order models with simple representational building blocks (e.g., chromatic pitch) to estimate and compare the average entropy of different musical works and corpora [26.16, 23–26] rather than dynamic prediction of ongoing sequential musical structure.

More recent research addressed these limitations by using variable-order Markov models, where the order varies depending on the context to generate accurate predictions dynamically throughout pieces of music [26.27, 28]. Prediction performance may also be improved by combining predictions from a long-term model (trained on a large corpus of music in the style) and a short-term model (which starts with an empty model and learns incrementally from the current musical work) [26.27, 28]. The long-term model represents the effects of long-term schematic exposure to a musical style while the short-term model reflects more local learning of repeated structure within a musical work. The probability distribution generated by the two models may be combined using arithmetic or geometric averaging, weighted by the entropy of the distributions [26.29].

Improved prediction performance can also be achieved by allowing the model to estimate and combine probabilities based on multiple features of the musical surface. Multiple viewpoint frameworks were originally developed by *Darrell Conklin* [26.27, 30, 31] to allow the integration of information from models of different features (inspired in part by *Ebcioğlu* [26.32]). The framework assumes a symbolic representation in which music is represented as a sequence of discrete events composed of a finite number of attributes each of which may assume a value from a finite alphabet. For example, a melody is often represented as a sequence of notes, each of which is composed of a pitch, onset time, duration and loudness.

A viewpoint is a mapping from a sequence of musical events to an element from the alphabet associated with the viewpoint. Basic viewpoints are projection functions associated with the attributes of the events (i. e., pitch, onset, duration and loudness in the example above). The framework also allows the specification of derived viewpoints (e.g., pitch interval, contour, scale degree), which are derived from a basic viewpoint (e.g., pitch in this case). Note that some viewpoints may be undefined at particular locations (e.g., pitch interval for the first event in a melody). Test viewpoints are viewpoints that return a Boolean value (e.g., whether a note falls on a tactus beat or not) and threaded viewpoints represent a base viewpoint (e.g., pitch interval) at points where a test viewpoint is true (e.g., pitch interval between notes falling on tactus beats) thereby allowing for sequences of nonsequential events. Finally, linked viewpoints represent the Cartesian product of two or more primitive viewpoints – for example, a link between pitch interval and scale degree will have an alphabet composed of pairs, whose first element is a pitch interval and whose second is a scale degree.

When modeling music with a multiple viewpoint system, separate models are constructed for each viewpoint included in the system and the resulting distributions are combined in much the same way as the long-term and short-term model outputs are combined. The distributions are first mapped back into distributions over the alphabet associated with the basic feature from which they are derived (e.g., pitch in the above examples) so that they can be combined. Typically, the viewpoint models are combined in a first stage separately for the long- and short-term models, which are then combined [26.27, 28, 33]. Multiple-viewpoint models have been developed for the domains of melody, harmony and voice leading [26.28, 33, 34].

When configured appropriately and trained on relevant corpora, these methods both improve prediction performance of the models [26.27, 28] and also account accurately for listeners' pitch expectations in melody [26.28, 35–40]. In some cases, the model parameters that optimize prediction performance do not improve fit to human perception and vice versa [26.28] suggesting that human expectation may be subject to constraints (such as memory or representational limitations) that prevent optimal prediction.

26.1.4 Beyond Simple Markov Models: Hidden Markov Models and Dynamic Bayesian Networks

The models of the Markov and n-gram family discussed in Sect. 26.1.3 are essentially equivalent to probabilistic versions of strictly local grammars (Chap. 25). One important feature of Markov models is that they can only model local sequential dependencies and do not assume any underlying deep structure (hidden variables). Although they operate directly on surface symbols, it is possible to use multiple viewpoint frameworks to allow such models to operate on nonsequential events (e.g., notes on tactus beats or phrase-final events, [26.27, 28]) and on representations of higher-order structure (e.g., phrase classes). However, although some of these formal limitations in expressive power may be addressed in part with the multiple-viewpoint approach, more expressive models of sequential structure have been developed in machine learning research, many of which have been applied to music. In the Chomsky hierarchy, the different model classes (from finite-state to finite-context) assume an underlying deep structure (represented using nonterminal symbols) that predicts the surface terminal symbols (Chap. 25); in an analogous way, many modeling approaches take advantage of an explicit representation of deep structure in music.

One well-known example is hidden Markov models (HMMs, e.g., [26.41, 42]; for an introduction see the comprehensive review by [26.43]), which correspond to probabilistic extensions of finite-state automata. As an extension of Markov models, the hidden Markov model assumes the Markov transition matrix not as a model characterizing transitions between surface symbols (as in the *visible* Markov models described above), but as transitions between deep structural (hidden) states that themselves emit surface symbols from associated emission distributions over the terminal alphabet. In other words, it assumes a Markov model as the underlying deep structure of states that govern the symbol distribution over subsections of the sequence.

HMMs have been employed to model various aspects of music, including, for instance, melodic structure [26.44, 45], meter and rhythm [26.45, 46], text setting [26.47] and harmonic structure [26.34, 48, 49]. Modeling harmony in a corpus of jazz standards, *Rohrmeier* and *Graepel* [26.34] found that a simple HMM modeling chord sequences exhibited barely any overfitting of its training data. Dynamic Bayesian networks (DBNs, [26.50]) generalize HMMs and constitute a family of graphical models that model the dependency structure of different (temporal) deep-structural features. DBNs were applied to modeling music by *Paiement* et al. [26.51] as well as *Raczynski* et al. [26.52] to model polyphonic pitch structure. Furthermore, DBNs can be straightforwardly adapted to modeling the interaction of different parallel feature-streams in the framework of HMMs. *Rohrmeier* and *Graepel* [26.34] implemented a DBN modeling Jazz harmony using features of duration and mode to improve predictive power though the approach does not extend to derived viewpoints.

Most of the models of sequence processing discussed so far drew their motivation from the modeling of musical expectation (see also [26.53], for an extensive account of the role of expectation in music perception). Models with a rich deep structure, however, may be used to understand other aspects of music processing such as inferring information encoded in the deep structure of the sequence. For instance *Raphael* and *Stoddard* [26.48] used a type of DBN for the inference of harmonic structure from the surface sequence of events. *Mavromatis* [26.45] employed a model selection procedure to find the optimal topology for a HMM, using this to draw theoretical and cognitive conclusions regarding representation of deep structure. He applied the model to two cases, the statistical learning and segmentation paradigm used by Saffran and colleagues (e.g., [26.54]) and metrical induction from rhythmical patterns in a corpus of Palestrina's vocal music. *Mavromatis* [26.46] extended this approach and drew computationally informed conclusions regarding a discussion surrounding Renaissance meter.

In summary, while the application of deep structure models and DBNs in cognitive music research is growing, there is a great potential to employ these models (in combination with model selection) to understand the role of deep structure in the representation and processing of musical syntax.

26.1.5 Hierarchical Models

While computational implementations of Markovian approaches (and derived approaches such as HMMs and DBNs) have largely addressed the problem of modeling effects of expectancy and prediction, many computational approaches have sought explicitly to implement hierarchical generative models of music. The motivation derives from theoretical insights that tonal music is organized in a hierarchical fashion and, accordingly, cognitive models of music processing should be able to account for such structural complexity. Moreover, theoretical hierarchical accounts of music stress that human music cognition involves substantially more than computation of sequential predictions including, in particular, the perception of large-scale processes (e.g., musical form), reductive listening, experience of hierarchical tension and recognizing similarity (see Chap. 25 for a discussion of various ways to motivate and explore the understanding of musical syntax). Accordingly, computational models of hierarchical structure have been inspired by a diverse array of modeling goals.

The hierarchical and generative branches of music theory mainly trace back to Schenkerian theory [26.55, 56]. Apart from Schenkerian theory itself, there have been three major lines of hierarchical modeling research: the generative theory of tonal music (GTTM) and tonal pitch space [26.57, 58], which originated from the goal of framing Schenkerian analysis in terms of formal linguistic approaches; *Narmour*'s theory of melodic expectancy and complexity [26.13, 59]; and approaches to tonal harmony that employ methods from generative linguistics and formal language theory [26.5, 6, 60, 61]. Each of these formal approaches have inspired efforts to build computational models.

Schenkerian Analysis and Derivative Models

Schenkerian theory constitutes one of the earliest and most comprehensive formal approaches to syntax in tonal music and remains dominant in music-theoretical teaching. It has been the object of several computational approaches from the early days of computing to the present day.

Early work by *Kassler* focused on understanding Schenkerian-like operations of analysis from the perspective of formal languages [26.62]. He formalized a subset of Schenkerian derivations with primitive-recursive functions [26.63] and described a basic implementation of a (presumably two-voice) model of Schenkerian analysis in terms of primitive operations on a matrix representation of pitch sequence, such as an Ursatz axiom, arpeggiation, neighbor note prolongation, simplified interruption (termed *articulation*), octave adjustment, bass ascent, mixture, etc. [26.62, 64, 65]. In a separate modeling attempt based on functional programming, Smoliar and colleagues used a recursive list structure of musical events that encoded a set of Schenkerian elaboration operations in terms of Lisp function calls. Drawing a direct formal analogy between (generative) linguistic parse trees and musical structure, they developed a tree-based structural representation of a Schenkerian reductive analysis [26.66–68]. In an approach similar to Schenkerian analysis techniques, *Baroni* and colleagues applied formal grammars to melodic structure [26.69, 70]. However, none of these early approaches resulted in a complete, fully automatic functioning model of Schenkerian analysis. *Marsden* [26.71] suggested that this may be due to to the massive explosion of combinatorial complexity that arises when encoding a formal account of Schenkerian-style reductive analysis at the level of notes.

With greater computational resources available, a number of researchers have recently returned to the problem of implementing Schenkerian analysis. *Mavromatis* and *Brown* [26.72] proposed a theoretical approach to modeling Schenkerian analysis with probabilistic context-free grammars. However, their model did not reach the stage of a full implementation due to issues of complexity (see [26.71]). *Marsden* [26.73] proposed a graph-theoretical representation of reductive structures (termed *E-Graph*) suitable for computational implementation. Subsequently, *Marsden* [26.74] proposed an expanded representation of Schenkerian reduction in a generative framework. Using the limited case of short musical phrases, Marsden developed a preliminary implementation of Schenkerian analysis [26.74] and building on this theoretical framework, a first proof-of-concept prototype was developed [26.71].

Recent research on implementing Schenkerian theory includes *Yust* [26.75], who proposed the structure of *maximal outerplanar graphs* (MOP) as representations for the structure of Schenkerian prolongation, and the work of *Kirlin* (e.g., [26.76]) who has developed a corpus of 41 pieces annotated with corresponding Schenkerian analyses in machine-readable format [26.77]. *Kirlin* and *Jensen* [26.78] use supervised probabilistic learning to uncover deep hierarchical musical structure, including an algorithm for deriving the most probable analysis for a given piece of music.

The Implication–Realization Model

Breaking with Schenkerian tradition (see [26.79]), *Eugene Narmour* developed a distinct theoretical approach to modeling melodic expectancy [26.13] as well as melodic complexity and reduction [26.59]. The *implication-realisation model* is based on the implications that a melodic interval (the implicative interval) has for the following interval (the realized interval) and presents a detailed classification of interval pairs based on the size and direction of the component intervals. The model consists of two independent perceptual systems – the bottom-up and top-down systems of melodic implication. While the principles of the former are held to be automatic, unconscious and universal, the principles of the latter are held to be learned and hence culture dependent. Although the model is presented in a highly analytic manner, it has psychological relevance because it proposes hypotheses about general perceptual principles that are precisely and quantitatively specified and therefore amenable to empirical investigation [26.80, 81]. Research has also compared the implication-realization model with variable-order Markov models in terms of how well they account for listeners' pitch expectations [26.28, 36, 37]. Furthermore, Grachten and colleagues have examined the use of the implication-realization model in a computational model of similarity for use in music information retrieval [26.82].

Furthermore, *Narmour* [26.59] presents detailed proposals for the ways in which basic musical structures (pairs of intervals) may form larger sequential units (chains) and larger hierarchical units (transformations) based on the idea of closure, which occurs when a structure does not generate a strong implication. However, these hierarchical aspects of the theory have been somewhat neglected in terms of computational modeling and empirical evaluation.

The Generative Theory of Tonal Music (GTTM)

The GTTM [26.57] constitutes probably the most influential model in empirical musicology to date. It models the perception of musical structure in term of the interaction of four levels of structure: grouping structure, metrical structure, time-span reduction and prolongational reduction (see [26.58, Ch. 1] for a review); these levels are defined in terms of *well-formedness rules* exhaustively defining the (very large) set of well-formed candidate analyses and *preference rules* that define rules to select the best analysis from the well-formed candidates. Although it is partly inspired by generative grammars for language, it is important to note that the GTTM is not a formal grammar because no (tree-defining) context-free rules are specified, and it is unclear whether it could be developed into one. Furthermore, the GTTM is a model of the final representation of a piece of Western tonal music as it might appear in the mind of an idealized listener, enculturated in Western music. It does not account for any effects of stylistic enculturation nor make any predictions about the dynamic nature of structure perception during listening (this was later addressed by *Jackendoff* [26.83]).

While the GTTM is specified to a much higher degree of detail and specificity than Schenkerian theory, it is still highly imprecise from a computational point of view. For instance, it does not specify a ranking for the preference rules and in some cases assumes human musical intuition for making analytical decisions. These factors pose challenges for computational implementations of the GTTM. Nonetheless, *Lerdahl* [26.58] has devised a model of musical tension that is based on a complete GTTM analysis and predicts musical tension based on local and global factors [26.58]. The model has been found to predict participants' continuous ratings of musical tension for a small number of musical pieces [26.84]. To date, the most extensive progress towards an implementation of the GTTM has been made by *Hamanaka* and colleagues [26.85, 86].

Generative Grammars for Music

Many authors have proposed recursive generative grammars for modeling the hierarchical organization of harmonic sequences, an idea whose essence goes back to *Riemann* [26.87]. Following the formalization of context-free rewrite grammars and the Chomsky hierarchy, a number of earlier approaches applied these techniques to music (e.g., [26.69, 70, 88, 89]). More recent approaches include *Steedman* [26.5, 6] and *Rohrmeier* [26.60, 61]. Based on Steedman's categorical grammar formalism [26.90], *Granroth-Wilding* and *Steedman* [26.91] implemented a model of jazz harmony that extends Steedman's earlier theoretical grammars for blues harmony [26.6]. Similarly, *De Haas* et al. developed an implementation of Rohrmeier's grammar for the purposes of music information retrieval, such as harmonic similarity and transcription [26.92–94]. *Tidhar* [26.95] implemented adapted versions of a context-free grammar formalism for the parsing of Couperin's unmeasured preludes.

Other approaches have attempted to combine context-free grammars with probabilistic learning. *Gilbert* and *Conklin* [26.96], for example, have employed a probabilistic context-free grammar for modeling melodic reduction. *Bod* [26.97] argues for a memory-based approach to modeling melodic grouping structure as an alternative to the Gestalt approach based on rules. He used grammar learning techniques to induce the annotated phrase structure of the Essen Folk Song Collection [26.98, 99]. Three grammar induction algorithms were examined: first, the treebank grammar learning technique, which reads all possible context-free rewrite rules from the training set and assigns each a probability proportional to its relative frequency in the training set; second, the Markov grammar technique, which assigns probabilities to context-free rules by decomposing the rule and its probability by an n-th-order Markov process, allowing the model to estimate the probability of rules that have not occurred in the training set; and third, a Markov grammar augmented with a data-oriented parsing (DOP) method for conditioning the probability of a rule over the rule occurring higher in the parse tree. A best-first parsing algorithm based on Viterbi optimization was used to generate the most probable parse for each melody in the test set given each of the three models. The results demonstrated that the treebank technique yielded moderately high precision but very low recall ($F = 0.065$), the Markov grammar yielded slightly lower precision but much higher recall ($F = 0.706$), while the Markov-DOP technique yielded the highest precision and recall ($F = 0.810$). A qualitative examination of the folk song data reveals a number of cases (15% of the phrase boundaries in the test set) where the annotated phrase boundary cannot be accounted for by Gestalt principles but that are predicted by the Markov-DOP parser.

26.1.6 Neural Networks

Neural networks represent a different class of models that have been used to understand the representation and processing of musical syntax. Rather than basing the model on a specific representation of musical structure (e.g., n-grams, production rules, music-theoretic principles), neural networks are biologically inspired in the sense that they take their motivation from basic properties of the brain (e.g., parallel processing across simple units, distributed representations, synaptic connectivity, graceful degradation and Hebbian learning). Note, however, that while they are *biologically inspired*, most neural network models (especially so-called connectionist models) are not actually *biologically plausible* models of neural processing. Thus they remain at *Marr*'s [26.100] algorithmic/representational level rather than at the implementational/physical level of description. At this level, one practical difficulty with neural networks as simulations of cognitive processing is that the nature of the learned representations are often not easily interpretable.

Mozer [26.101], for instance, developed a model based on a recurrent artificial neural network (RANN, [26.102]) and used psychoacoustic constraints in the representation of pitch and duration. In particular, the networks operated over predefined, theoretically motivated multidimensional spatial representations of pitch (which emphasized a number of pitch relations including pitch height, pitch chroma and fifth relatedness, [26.103]) and duration (emphasizing such relations as relative duration and tuplet class). These neural networks are trained within a supervised regime in which the discrepancy between the activation of the output units (the expected next event) and the desired activation (the actual next event) is used to adjust the network weights at each stage of training. When trained and tested on sets of simple artificial pitch sequences with a split-sample experimental paradigm, the RANN model outperformed simple digram models. In particular, the use of cognitively motivated multidimensional spatial representations led to significant benefits (over a local pitch representation) in the training of the networks. However, the results were less than satisfactory when the model was trained on a set of melodic lines from ten compositions by J.S. Bach and used to generate new melodies; the neural network architecture appeared unable to capture the higher-level structure in these longer pieces of music.

The question arises of whether these neural network models provide good simulations of human processing. In an artificial grammar learning paradigm of melodic structure [26.40, 104], a simple recurrent network model [26.102] was compared with an n-gram model [26.28] and a chunking model [26.105]. Results indicate that the n-gram model achieved by far the best performance, yet the simple recurrent network exhibited characteristic patterns of performance (including errors) that were closer to the human level [26.106, 107].

One approach to addressing the apparent inability of RANNs to represent recursive constituent structure in music involves what is called auto-association. An auto-associative network is simply one that is trained to reproduce on its output layer a pattern presented to its input layer, generally forming a compressed representation of the input on its hidden layer. For example, training a network with eight-unit input and output layers separated by a three-unit hidden layer with the eight one-bit-in-eight patterns typically results in a three-bit binary code on the hidden units [26.108]. *Pollack* [26.109] introduced an extension of auto-association called recursive auto-associative memory (RAAM), which is capable of learning fixed-width representations for compositional tree structures through repeated compression. The RAAM architecture consists of two separate networks: first, an encoder network that constructs a fixed-dimensional code by recursively processing the nodes of a symbolic tree from the bottom up; and second, a decoder network that recursively decompresses this code into its component parts until it terminates in symbols, thus reconstructing the tree from the top down. The two networks are trained in tandem as a single auto-associator. *Large* et al. [26.110] examined the ability of RAAM to acquire reduced representations of Western children's melodies represented as tree structures according to music-theoretic predictions [26.57]. It was found that the trained models acquired compressed representations of the melodies in which structurally salient events are represented more efficiently (and reproduced more accurately) than other events. Furthermore, the trained network showed some ability to generalize beyond the training examples to variant and novel melodies although, in general, performance was affected by the depth of the tree structure used to represent the input melodies with greater degrees of hierarchical nesting leading to impaired reproduction of input melodies. However, the certainty with which the trained network reconstructed events correlated well with music-theoretic predictions of structural importance [26.57] and cognitive representations of structural importance as assessed by empirical data on the events retained by trained pianists across improvised variations on the melodies.

Recent developments in neural networks have led to successful modeling of musical structure using restricted Boltzmann machines (RBMs) [26.111, 112]. RBMs appear to be approaching the prediction performance of the best-performing variable-order Markov models described in Sect. 26.1.3 [26.112]. However, these neural network models are difficult to analyze and it remains to be seen how successful they will be in modeling cognitive processing of musical syntax. Nonetheless, they do at least demonstrate that such processing can be implemented using parallel distributed (and potentially nonsymbolic) representations.

26.2 Psychological Research

Computational models shed light on the plausibility of specific assumptions about, and constraints on, the nature of the cognitive representations and algorithms underlying music perception and serve as analytical tools to explore the types of syntactic structures present in music. However, to understand the kinds of syntactic structures that are perceived and represented by listeners, it is important to compare empirically the output of a model with the responses of listeners. Empirical research on musical listening can help identify the power, the limits and constraints of human perception and cognition of musical syntax. In this context, computational models are useful for generating hypotheses and selecting stimuli that differ quantitatively in terms of their syntactic properties (e.g., grammaticality, uniformity, see Chap. 25). Although there are notable differences between language and music (e.g., in terms of lexical categories and the nature of semantic content), cognitive scientific research on music follows an analogous approach, exploring processing via a combination of theoretical inquiry, computational modeling and psychological/neuroscientific testing [26.113]. Furthermore it has been suggested that music and language, which are both sequential auditory forms of human communication, may share similarities at more abstract levels of cognitive and neural representation [26.114–116], a point to which we return below.

Psychological research has established that enculturated listeners are sensitive to sequential structure in music; however, research paradigms have sometimes been driven by (overly) simple representational assumptions. We will examine the evidence relating to harmonic movement, melody and high-level form. One topic that continues to attract debate is the extent to which listeners are capable of representing hierarchical, nonsequential relationships in music. We examine research on this question below.

26.2.1 Perception of Local Dependencies

As discussed in Chap. 25, harmony constitutes one of the core building blocks in Western tonal music and it has been studied in the context of different theories of syntax. Accordingly, a large number of experimental studies have focused on the perception of harmonic structure from a variety of perspectives.

At a local level, listeners are able to detect harmonic relations between successive chords as evidenced by longer reaction times to in-key than out-of-key harmonic transitions in the paradigm of musical *priming* studies ([26.117, 118], see [26.119] for a review). These findings have been replicated with longer sequences of chords and more complex harmonic relations [26.120, 121]. Furthermore, there is evidence that these harmonic priming effects are affected both by the global context of the prevailing key (which determines the overall stability of a chord), by the local harmonic context and by the rhythmic organization of the progression [26.120]. Research has also examined whether these effects are best explained by learned properties of harmonic movement or by sensory properties of the target chords [26.121–123]. The results consistently show that the structural properties of harmonic movement have a stronger effect than sensory influences such as repetition priming [26.124, 125], even when efforts are made to make these sensory influences very strong. Furthermore, *Tillmann* et al. [26.126] used self-organizing maps (a variety of neural network) to argue that these priming effects can be explained by learning of tonal organization through musical exposure.

We can ask similar questions about the perception of structure in melody. Early work on predictive processing of musical structure focused on rules of melodic organization and how they influence pitch expectations [26.13, 80, 81]. Empirical studies of these rules have found that listeners' melodic expectations do generally exhibit influences of pitch proximity (smaller intervals are more expected) and pitch reversal (large pitch intervals are expected to be followed by smaller ones in the opposite registral direction) [26.53, 127]. Actual melodies also exhibit these properties [26.128], possibly reflecting physical constraints of performance – the difficulty of producing large intervals accurately and tessitura constraints producing regression to the mean after large intervals [26.129–131]. Therefore, it remains possible that listeners learn these regularities (or variants of them, subject to cognitive constraints, [26.53, 132]), which are then reflected in their pitch expectations.

In fact, there is empirical evidence of implicit learning of regularities in musical melody and other sequences of pitched events ([26.40, 54, 104, 133]; see [26.8] for a review on implicit learning of music). Consistent with an approach based on statistical learning, melodic pitch expectations vary between musical styles [26.134] and cultures [26.135–140], throughout development [26.141] and across degrees of musical training and familiarity [26.36, 37, 134]. Furthermore, pitch expectations appear to be informed both by long-term exposure to music [26.142] and by the encoding of regularities in the immediate context [26.133]. As with harmonic expectations, computational models relying on nonhierarchical statistical learning of regularities have proved highly successful in predicting listeners' melodic expectations [26.28, 37, 38].

All of these effects can be accounted for by local models corresponding to strictly local grammars. Therefore, it is crucial to ask which aspects of hierarchical and nonlocal structures affect music listening and processing.

26.2.2 Perception of Nonlocal Dependencies

There are several ways in which nonlocal dependencies can be expressed in music, ranging from short chord sequences (such as "I IV V/ii ii V I", in which the IV chord implies the V chord and not V/ii, and the second I chord prolongs the initial one, rendering the whole sequence a constituent of a prolonged tonic) to dependencies between (sub)phrases (e.g., antecedent-consequent patterns) and between larger formal units like the parts of a minuet or a sonata [26.143]. It may even be possible to understand these dependencies at different timescales (chords, parts of phrases, phrases, movements) as recursive instantiations of similar structures [26.57, 143].

Deutsch [26.144] describes experiments in which subjects were presented with sequences of 12 notes, which they recalled in musical notation. Half the sequences were structured in accordance with the model of *Deutsch* and *Feroe* [26.10], described above, such that a higher-level subsequence of four elements acted on a lower-level subsequence of three or four elements while the remaining sequences were unstructured. Sequences were presented in one of three conditions: first, with no temporal segmentation; second, temporally segmented in accordance with tonal structure; and third, temporally segmented in violation of tonal structure. The results demonstrated that recall was high in the first and second conditions and low in the third condition and for unstructured sequences, suggesting that hierarchically structured sequences are better encoded in memory. However, on a similar task, *Boltz* and *Jones* [26.145] have found that rule recursion has only a modest effect on memory for melodies and only in certain conditions.

Regarding large-scale form, researchers have studied the aesthetic judgments of musicians and nonmusicians for pieces of music in which the large-scale tonal form has been disrupted by rearranging or rewriting certain parts. These studies have been conducted by rearranging the various movements of Beethoven sonatas and string quartets [26.146], reordering the variations making up Bach's Goldberg variations [26.147], altering the endings of excerpts from Romantic and Classical works (1–6 min) such that they start and end in a different key ([26.148]; methodologically criticized by *Gjerdingen* [26.149]) and using rearranged versions of the opening movement of Mozart's Symphony in G Minor [26.150]. The results consistently suggest that listeners gain as much pleasure from the altered versions as the unaltered versions and musicians are no more affected than nonmusicians by these disruptions [26.148, 150]. In a study of musicians listening to original and rearranged versions of six keyboard works by Handel [26.151] found that both versions were deemed equally conforming to stylistic expectations and accuracy in judging whether the starting and ending key were the same was at chance. These findings suggest that there are severe constraints on the ability to build cognitive representations of large-scale formal relationships in music, even for musicians (though none of these studies examined professional musicians with a very high level of expertise).

However, there is crucial evidence that listeners are sensitive to long-distance hierarchical dependencies at the phrase-level and midrange timescales. *Koelsch* et al. [26.152], examined responses to the final chord in a pair of chorale melodies composed of two subphrases, together with modified versions in which the first phrase was altered to break tonal closure with the final phrase. The results showed strong characteristic differences between the two versions in the neural response (the early right anterior negativity (ERAN), Sect. 26.3.2) to the final chord. In addition, behavioral measures showed that there were no differences in emotional response (valence and arousal) between the versions, suggesting that nonlocal harmonic dependencies are independent of emotional expression. Further behavioral measures showed that the final chord of the original version was judged to close the sequence better than the final chord of the altered versions (although the difference was relatively small, pointing towards the use of implicit rather than explicit knowledge). Fundamentally, because the harmonic dependencies in this study exceed ten chords, it is impossible that nonveridical *n*-gram or Markov models could explicitly represent the difference. Accordingly, this evidence for nonlocal dependencies affecting the perception of phrase closure falsifies the assumption that simple local models of harmony can adequately model human perception of musical syntax.

26.3 Neuroscientific Research

26.3.1 Introduction

It is pertinent to ask what cognitive neuroscience can tell us over and above research in experimental psychology and cognitive science [26.153–156]. Clearly, it can tell us something about the neural basis of psychological processes; less obviously, but perhaps more importantly, it can also tell us something about those psychological processes themselves. Once a neural response has been linked to a specific psychological process (and this itself may require great effort to establish), then we can use it as an additional measure of that process, alongside behavioral measures. Such neural responses have the potential advantages that they may be more sensitive and less prone to various kinds of bias than behavioral measures. Potential disadvantages include the difficulty of establishing direct, specific relationships between features of the neural response and particular psychological processes.

26.3.2 Neural Basis of Syntactic Processing in Music

Neuroscientific research has used electroencephalography (EEG) to investigate event-related-potential (ERP) responses to violations of harmonic syntax [26.157–164]. Two characteristic brain responses have been reported: an early anterior negativity (EAN) with a latency of 150–280 ms (sometimes right lateralized and referred to as the ERAN – early right anterior negativity), and a later bilateral or right-lateralized negativity (N5) with a latency of 500 ms [26.157, 161]. The EAN is thought to reflect the violation of harmonic expectation, while the N5 is thought to reflect the higher processing effort needed to integrate unexpected harmonies into the ongoing context [26.163]. The amplitude of the EAN is related to the long-term transition probability of the chord [26.165, 166] suggesting that it reflects implicit learning of harmonic movement through experience (see Sect. 26.1.3 above). Consistent with this proposal are findings that the EAN is attenuated (though still present) in five to six-year-old children compared to adults and accentuated in adult musicians, relative to adult nonmusicians [26.159].

To date, less is known about the neural correlates of structural processing in other aspects of music such as pitch, rhythm and timbre. Early studies [26.167–171] identified a late positive component (LPC) peaking between 300–600 ms at central and posterior sites in response to stylistically unexpected notes in a melody.

The amplitude and latency of the LPC are sensitive to musical expertise, the familiarity of the melody, the degree of expectancy violation [26.167], and also to the timing of the unexpected note [26.168].

There is an important distinction between schematic representations of the syntactic rules governing a musical style and veridical memory for the structure of a familiar piece of music [26.172]. *Miranda* and *Ullman* [26.173] describe a functional dissociation between two ERP components: an early (150–270 ms) anterior-central negativity (i. e., the EAN) associated with out-of-key violations in both familiar and unfamiliar melodies, and a subsequent (220–380 ms) posterior negativity elicited by both in-key and out-of-key violations of familiar melodies only. They suggested that these two components are driven by violations of musical rules (of tonality/harmony) and of veridical memory representations of familiar melodies respectively. In order to focus exclusively on schematic acquisition of syntax, *Loui* et al. [26.166] examined neural responses to deviant melodic endings in pitch sequences generated according to an artificial (strictly local) grammar, using pitches taken from the unfamiliar Bohlen–Pierce scale. They report an EAN, whose amplitude increased with greater learning of the grammar (measured by degree of exposure and performance in a grammaticality decision task). Again, this suggests that the EAN reflects a process of implicit statistical learning of sequential dependencies in the auditory environment (Sect. 26.1.3).

The EAN to violations of melodic syntax tends to occur earlier (around 100 ms) than to violations of harmonic syntax (circa 180 ms) [26.37, 38, 174], perhaps indicating that single notes are processed more quickly than chords. Using a computational model of auditory expectation [26.28], which makes probabilistic pitch predictions based on statistical learning, to identify notes in melodies that varied systematically in information content (IC), *Omigie* et al. [26.39] showed a linear relationship between the amplitude of the EAN and IC. *Pearce* et al. [26.37] reported an increase in the amplitude of beta oscillations for high information content (low probability) notes at a latency of around 500 ms in centroparietal regions and in phase locking in the same time window between electrodes located over centroparietal and occipital regions. Again these results are consistent with the proposal that these neural indicators of syntactic processing reflect probabilistic inference based on implicit statistical learning.

26.3.3 Syntax in Music and Language

EEG research has also addressed the question of whether there exists an overlap in neural processing of musical and linguistic syntax. Violations of syntax in language often generate two characteristic ERPs: a left anterior negativity (LAN) peaking around 300–450 ms at frontal scalp locations, and a positive-going deflection termed the P600 at a latency of about 600 ms with a posterior distribution. An early study suggested that violations of harmonic syntax generate an increased P600, which is very similar to that induced by violations of linguistic syntax [26.162]. More recent research has presented music synchronously with visually presented sentences where each word coincides temporally with a note or chord in the music. Introducing syntactic violations in the music and language allows the investigation of conditions where the violations are congruent or incongruent in the two domains. This research suggests that the LAN to syntactic violations in language is reduced by unusual harmonic movement (e.g., a Neapolitan chord, [26.175]) and also by low-probability notes in melodic phrases [26.176]. There is also evidence that the ERAN is reduced when presented concurrently with a linguistic violation [26.177]. Interestingly, these interactive effects are not in evidence when musical violations are paired with semantic incongruities in language [26.176, 177].

Analogies have been drawn between the ERAN and the early left anterior negativity (ELAN) often observed in response to violations of syntax in language [26.178]. Interestingly, children with specific language impairment not only show characteristic changes in the ELAN to language [26.179] but also a reduced ERAN to harmonic violations [26.180]. Broca's aphasics also show reduced neural responses to harmonic violations in music [26.181] and there is evidence from magnetoencephalography (MEG) research that the ERAN to violations of harmonic syntax originates in Broca's area and its right hemisphere homologue [26.182]. Furthermore, functional magnetic resonance imaging (fMRI) studies [26.183–186] suggest that violations of harmonic syntax induce activation in the inferior frontal cortex, which is also suggestive of a relationship between the neural processing of syntax in music and language [26.114].

26.3.4 Grouping Structure

One other aspect of musical structure that has been studied from the perspective of cognitive neuroscience is grouping structure. Using EEG and MEG, *Knösche* et al. [26.187] found that phrase boundaries in melodies generated a late (500–600 ms for EEG, 400–700 ms for MEG) positive deflection, which they termed the *closure positive shift (CPS)*. Source localization suggested that the CPS was generated by structures in the limbic system, including anterior/posterior cingulate and posterior mediotemporal cortex. Similar neural responses have been observed at syntactic phrase boundaries in language [26.188], which seem to be related to prosodic cues [26.189, 190]. A subsequent study showed that the CPS to musical phrase boundaries is stronger in musicians than in nonmusicians [26.191], suggesting that strategies for segmenting music are influenced by musical training.

26.4 Implications and Issues

We have reviewed empirical research on musical syntax from computational, psychological and neuroscientific perspectives. We close with a discussion of two issues that naturally arise during this discussion: first, the extent to which the different perspectives on musical syntax converge; and second, the relationship between syntax and semantics in music, which naturally invokes the question of affective responses to music.

26.4.1 Convergence Between Approaches

One of the key challenges facing future research is to integrate insights from these different methodological and epistemological approaches. Computational modeling allows us to implement theories of musical structure and syntax and examine the behavior of the implemented algorithm when supplied with musical examples. This can provide insights into optimal syntactic representations and absolute constraints on the syntactic structure of a musical style [26.192]. Psychological research, on the other hand, can indicate the kinds of syntactic forms and relationships that listeners are capable of perceiving, representing and learning. Implementing a psychological theory as a computational model requires the theory to be precisely expressed and also allows the theory to be tested and refined through quantitative comparison of human and model responses. It also allows comparison of human performance at different levels of experience and expertise with optimal syntactic parsing. Finally, neuroscientific research provides information about the neural basis of syntactic processing, which can impose constraints on human syntactic processing and also provide data that is more sensitive to implicit knowledge than behavioral methods (e.g., [26.193, 194]. Therefore, future research should triangulate more explicitly between computational modeling, psychological experimentation and cognitive-neuroscientific investigation in further developing our understanding of musical syntax [26.37, 195].

26.4.2 Syntax, Semantics and Emotion

Musical styles can be said to possess syntax in an analogous way to that in which natural languages (or programming languages) do. However, music is different from natural language in that musical elements do not usually carry clear referential and propositional semantics in the way that linguistic atoms do. It is sometimes possible to establish indexical references for appropriately enculturated listeners – the old castle, the sea, a storm, the spring, love, James Bond's theme or Brunhilde's leitmotif being good examples taken from various pieces of music. However, it is impossible to communicate complex statements like *had he not gone to sea last spring and then returned to the old castle, James Bond would not have fallen in love with Brunhilde* (see [26.113, 196, 197] for further discussion). Therefore, meaning in musical communication is borne to a large extent by syntactic structure and the listener's perception of structural relations between musical elements [26.198]. In particular, the syntactic structure of music affords the communication of patterns of tension and resolution, through the systematic manipulation of the listener's structural expectations, based in turn on their internalized syntactic representations of the style.

There is one prominent theoretical perspective on affective responses to music that is relevant here since it relates the expression and perception of meaning to predictive processing of musical structure, using information-theoretic principles [26.53]. Building on arguments made by *Hanslick* [26.199], *Meyer* [26.198, 200] examines from a theoretical perspective the dynamic cognitive processes in operation when we listen to music and how these processes not only underlie the listener's understanding of musical structure but also give rise to the communication of affect and the perception of meaning in music. Meyer proposes that meaning arises through the manner in which musical structures activate, inhibit and resolve expectations in the listener about forthcoming musical structures. He notes that these expectations may differ independently in terms of the degree to which they are passive or active, their strength and their specificity. He contends, in particular, that affect is aroused when a passive expectation induced by antecedent musical structures is made active by it being temporarily inhibited or permanently blocked by consequent musical structures. Meyer discusses three ways in which the listener's expectations may be violated. The first occurs when the expected consequent event is delayed, the second when the antecedent context generates ambiguous expectations about consequent events, and the third when the consequent is unexpected. While the particular effect of music is clearly dependent on the strength of the expectation, Meyer argues that it is also conditioned by the specificity of the expectation.

Meyer [26.200] discusses the relationship between his theory of musical expectancy and concepts in information theory. He starts with the suggestion that [26.200, p. 414]:

> *once a musical style has become part of the habit responses of composers, performers and practiced listeners it may be regarded as a complex system of probabilities*

and that expectations arise out of these internalized probability systems. In particular, he suggests that a musical style may be conceived as a Markov process (Sect. 26.1.3) and that experienced listeners possess internalized models of that process. The degree to which hypothetical meanings provide ambiguous expectations about consequent structures can be measured by entropy (or uncertainty) [26.200, p. 416]:

> *The lower the probability of a particular consequent* [...] *the greater the uncertainty (and information) involved in the antecedent-consequent relation.*

An unexpected consequent conveys a maximum of information. The process of revaluation corresponds to the feedback of information such that future behavior is conditioned by the results of past events.

Meyer notes that uncertainty may arise from different sources. Thus, systemic uncertainty decreases throughout the experience of a piece of music as the listener's model develops and the composer may deliberately introduce designed uncertainty to combat this effect. Furthermore, the redundancy (lack of uncertainty) inherent in a style serves to combat noise, be it cultural (resulting from discrepancies between the habit responses of a given listener and those operating in the style) or acoustical. *Witten* et al. [26.201, p.71] make a similar distinction between perceptual uncertainty (that which is relative to a particular listener's model) and stylistic uncertainty (that which is inherent in the musical style).

As noted above, the most obvious emotional response to expectancy violation and uncertainty in music is tension but according to *Meyer* [26.198], these processes may also give rise to a range of specific emotional experiences including apprehension/anxiety (p.27), hope (p.29), and disappointment (p.182) (see also [26.202]). Convergingly, empirical research has found that stylistically unusual chord progressions do stimulate increases in physiological arousal [26.177] while notes that have a low probability of occurrence in performed melodic music [26.28] have been found to be associated with increased arousal, reduced valence, increased skin conductance and reduction in heart rate [26.35].

References

26.1 P.N. Johnson-Laird: *Mental Models* (Harvard Univ. Press, Cambridge 1983)

26.2 H.C. Longuet-Higgins: Artificial intelligence – A new theoretical psychology?, Cognition **10**(1–3), 197–200 (1981)

26.3 H.A. Simon, C.A. Kaplan: Foundations of cognitive science. In: *Foundations of Cognitive Science*, ed. by M.I. Posner (MIT Press, Cambridge 1989) pp. 1–47

26.4 A. Newell, H.A. Simon: Computer science as empirical enquiry: Symbols and search, Commun. ACM **19**(3), 113–126 (1976)

26.5 M. Steedman: A generative grammar for jazz chord sequences, Music Percept. **2**(1), 52–77 (1984)

26.6 M. Steedman: The blues and the abstract truth: Music and mental models. In: *Mental Models in Cognitive Science*, ed. by A. Garnham, J. Oakhill (Erlbaum, Mahwah 1996) pp. 305–318

26.7 A. Clark: Learning trees from strings: A strong learning algorithm for some context-free grammars, J. Mach. Learn. Res. **14**, 3537–3559 (2013)

26.8 M. Rohrmeier, P. Rebuschat: Implicit learning and acquisition of music, Top. Cogn. Sci. **4**(4), 525–553 (2012)

26.9 N. Chomsky: *Aspects of the Theory of Syntax* (MIT Press, Cambridge 1965)

26.10 D. Deutsch, J. Feroe: The internal representation of pitch sequences in tonal music, Psychol. Rev. **88**(6), 503–522 (1981)

26.11 H.A. Simon, R.K. Sumner: Pattern in music. In: *Formal Representation of Human Judgement*, ed. by B. Kleinmuntz (Wiley, New York 1968) pp. 219–250

26.12 L.B. Meyer: *Explaining Music: Essays and Explorations* (University of Chicago Press, Chicago 1973)

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

26.14 C. Ames: Automated composition in retrospect: 1956–1986, Leonardo **20**(2), 169–185 (1987)

26.15 C. Ames: The Markov process as a compositional model: A survey and tutorial, Leonardo **22**(2), 175–187 (1989)

26.16 J.E. Cohen: Information theory and music, Behav. Sci. **7**(2), 137–163 (1962)
26.17 L. Hiller: Music composed with computers – A historical survey. In: *The Computer and Music*, ed. by H.B. Lincoln (Cornell Univ. Press, Cornell 1970) pp. 42–96
26.18 C.E. Shannon: A mathematical theory of communication, Bell Syst. Tech. J. **27**(3), 379–423 (1948), and 623–656
26.19 M.A. Rohrmeier, S. Koelsch: Predictive information processing in music cognition. A critical review, Int. J. Psychophysiol. **83**(2), 164–175 (2012)
26.20 M.T. Pearce, G.A. Wiggins: Auditory expectation: The information dynamics of music perception and cognition, TopiCS Cogn. Sci. **4**, 625–652 (2012)
26.21 D.J.C. MacKay: *Information Theory, Inference, and Learning Algorithms* (Cambridge Univ. Press, Cambridge 2003)
26.22 L. Hiller, L. Isaacson: *Experimental Music* (McGraw–Hill, New York 1959)
26.23 L. Hiller, C. Bean: Information theory analyses of four sonata expositions, J. Music Theory **10**(1), 96–137 (1966)
26.24 L. Hiller, R. Fuller: Structure and information in Webern's Symphonie, Op. 21, J. Music Theory **11**(1), 60–115 (1967)
26.25 R.C. Pinkerton: Information theory and melody, Sci. Am. **194**(2), 77–86 (1956)
26.26 J.E. Youngblood: Style as information, J. Music Theory **2**, 24–35 (1958)
26.27 D. Conklin, I.H. Witten: Multiple viewpoint systems for music prediction, J. New Music Res. **24**(1), 51–73 (1995)
26.28 M.T. Pearce: *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition*, Ph.D. Thesis (London City University, London 2005)
26.29 M.T. Pearce, D. Conklin, G.A. Wiggins: Methods for combining statistical models of music. In: *Computer Music Modelling and Retrieval*, ed. by U.K. Wiil (Springer, Berlin, Heidelberg 2005) pp. 295–312
26.30 D. Conklin: *Prediction and Entropy of Music*, Master's dissertation (University of Calgary, Calgary 1990)
26.31 D. Conklin, J.G. Cleary: Modelling and generating music using multiple viewpoints. In: *Proc. 1st Workshop AI Music, Menlo Park* (1988) pp. 125–137
26.32 K. Ebcioğlu: An expert system for harmonising four-part chorales, Comput. Music J. **12**(3), 43–51 (1988)
26.33 R. Whorley, G. Wiggins, C. Rhodes, M.T. Pearce: Multiple viewpoint systems: Time complexity and the construction of domains for complex musical viewpoints in the harmonisation problem, J. New Music Res. **42**, 237–266 (2013)
26.34 M. Rohrmeier, T. Graepel: Comparing feature-based models of harmony. In: *Proc. 9th Int. Symp. Comput. Music Model. Retr.* (Springer, London 2012) pp. 357–370
26.35 H. Egermann, M.T. Pearce, G.A. Wiggins: McAdams: Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music, Cogn. Affect. Behav. Neurosci. **13**, 533–553 (2013)
26.36 N.C. Hansen, M.T. Pearce: Predictive uncertainty in auditory sequence processing, Front. Psychol. **5**, 1052 (2014), https://doi.org/10.3389/fpsyg.2014.01052
26.37 M.T. Pearce, M.H. Ruiz, S. Kapasi, G.A. Wiggins, J. Bhattacharya: Unsupervised statistical learning underpins computational, behavioural and neural manifestations of musical expectation, NeuroImage **50**, 302–313 (2010)
26.38 D. Omigie, M.T. Pearce, L. Stewart: Tracking of pitch probabilities in congenital amusia, Neuropsychologia **50**, 1483–1493 (2012)
26.39 D. Omigie, M.T. Pearce, V. Williamson, L. Stewart: Electrophysiological correlates of melodic processing in congenital amusia, Neuropsychologia **51**, 1749–1762 (2013)
26.40 M. Rohrmeier, I. Cross: Artificial grammar learning of melody is constrained by melodic inconsistency: Narmour's principles affect melodic learning, PLOS ONE **8**(7), e66174 (2013)
26.41 L.E. Baum, T. Petrie: Statistical inference for probabilistic functions of finite state Markov chains, Ann. Math. Stat. **37**(6), 1554–1563 (1966)
26.42 L.E. Baum, T. Petrie, G. Soules, N. Weiss: A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, Ann. Math. Stat. **41**(1), 164–171 (1970)
26.43 L.R. Rabiner: A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE **77**(2), 257–286 (1989)
26.44 P. Mavromatis: A hidden Markov model of melody production in greek church chant, Comput. Musicol. **14**, 93–112 (2005)
26.45 P. Mavromatis: HMM analysis of musical structure: Identification of hidden variables through topology-sensitive model selection. In: *Mathematics and Computation in Music*, Communications in Computer and Information Science, (Springer, Berlin, Heidelberg 2009) pp. 205–217
26.46 P. Mavromatis: Exploring the rhythm of the Palestrina style. A case study in probabilistic grammar induction, J. Music Theory **56**(2), 169–223 (2012)
26.47 P. Mavromatis: Minimum description length modeling of musical structure, J. Math. Music **3**(3), 117–136 (2009)
26.48 C. Raphael, J. Stoddard: Functional harmonic analysis using probabilistic models, Comput. Music J. **28**(3), 45–52 (2004)
26.49 J.P. Bello, J. Pickens: A robust mid-level representation for harmonic content in music signals. In: *ISMIR*, Vol. 5 (2005) pp. 304–311
26.50 K.P. Murphy: *Dynamic Bayesian Networks: Representation, Inference and Learning*, Doctoral Dissertation (Univ. of California, Berkeley 2002)
26.51 J.-F. Paiement, Y. Grandvalet, S. Bengio: Predictive models for music, Connect. Sci. **21**, 253–272 (2009)

26.52 S.A. Raczyński, S. Fukayama, E. Vincent: Melody harmonization with interpolated probabilistic models, J. New Music Res. **42**(3), 223–235 (2013)
26.53 D. Huron: *Sweet Anticipation: Music and the Psychology of Expectation* (MIT Press, Cambridge 2006)
26.54 J.R. Saffran, E.K. Johnson, R.N. Aslin, E.L. Newport: Statistical learning of tone sequences by human infants and adults, Cognition **70**(1), 27–52 (1999)
26.55 H. Schenker: *Der Freie Satz. Neue Musikalische Theorien und Phantasien* (Margada, Liège 1935)
26.56 A. Forte, S.E. Gilbert: *Introduction to Schenkerian Analysis* (Norton, New York 1982)
26.57 F. Lerdahl, R. Jackendoff: *A Generative Theory of Tonal Music* (MIT Press, Cambridge 1983)
26.58 F. Lerdahl: *Tonal Pitch Space* (Oxford Univ. Press, Oxford 2001)
26.59 E. Narmour: *The Analysis and Cognition of Melodic Complexity: The Implication-realisation Model* (University of Chicago Press, Chicago 1992)
26.60 M. Rohrmeier: A generative grammar approach to diatonic harmonic structure. In: *Proc. 4th Sound Music Comput. Conf. (SMC07)*, ed. by C. Spyridis, A. Georgaki, G. Kouroupetroglou, C. Anagnostopoulou (2007) pp. 97–100
26.61 M. Rohrmeier: Towards a generative syntax of tonal harmony, J. Math. Music **5**(1), 35–53 (2011)
26.62 M. Kassler: Proving musical theorems I: The middleground of Hienrich Schenker's theory of tonality, Technical Report no. 103, (University of Sydney, Sydney 1975)
26.63 M. Kassler: *A Trinity of Essays*, Ph.D. Thesis (Princeton University, Princeton 1967)
26.64 M. Kassler: Explication of the middleground of Schenker's theory of tonality, Misc. Musicol. Adel. Stud. Music. **9**, 72–81 (1977)
26.65 M. Kassler: APL applied in music theory, APL Quote Quad **18**, 209–214 (1988)
26.66 R.E. Frankel, S.J. Rosenschein, S.W. Smoliar: Schenker's theory of tonal music – its explication through computational processes, Int. J. Man-Mach. Stud. **10**, 121–138 (1976)
26.67 R.E. Frankel, S.J. Rosenschein, S.W. Smoliar: A LISP-based system for the study of Schenkerian analysis, Comput. Humanit. **10**, 21–32 (1978)
26.68 S.W. Smoliar: A computer aid for Schenkerian analysis, Comput. Music J. **4**, 41–59 (1980)
26.69 M. Baroni: The concept of musical grammar (translated by S. Maguire with the assistance of W. Drabkin), Music Anal. **2**, 175–208 (1983)
26.70 M. Baroni, R. Dalmonte, C. Jacobini: Theory and analysis of European melody. In: *Computer Representations and Models in Music*, ed. by A. Marsden, A. Pople (Academic Press, London 1992) pp. 187–206
26.71 A. Marsden: Schenkerian analysis by computer: A proof of concept, J. New Music Res. **39**, 269–289 (2010)
26.72 P. Mavromatis, M. Brown: Parsing context-free grammars for music: A computational model of Schenkerian analysis. In: *Proc. 8th Int. Conf. Music Percept. Cogn., Evanston* (2004) pp. 414–415
26.73 A. Marsden: Representing melodic patterns as networks of elaborations, Comput. Humanit. **35**, 37–54 (2001)
26.74 A. Marsden: Generative structural representation of tonal music, J. New Music Res. **34**, 409–428 (2005)
26.75 J.D. Yust: *Formal Models of Prolongation*, Ph.D. Thesis (University of Washington, Washington 2006)
26.76 P.B. Kirlin: Using harmonic and melodic analyses to automate the initial stages of Schenkerian analysis. In: *Proc. Int. Conf. Music Inf. Retr. (ISMIR), Kobe* (2009) pp. 423–428
26.77 P.B. Kirlin: A data set for computational studies of Schenkerian analysis. In: *Proc. 15th Int. Soc. Music Inf. Retr. Conf* (2014) pp. 213–218
26.78 P.B. Kirlin, D.D. Jensen: Using supervised learning to uncover deep musical structure. In: *Proc. 29th AAAI Conf. Artif. Intell.* (2015) pp. 1770–1776
26.79 E. Narmour: *Beyond Schenkerism: The Need for Alternatives in Music Analysis* (University of Chicago Press, Chicago 1975)
26.80 C.L. Krumhansl: Music psychology and music theory: Problems and prospects, Music Theory Spectr. **17**, 53–90 (1995)
26.81 E.G. Schellenberg: Expectancy in melody: Tests of the implication-realisation model, Cognition **58**(1), 75–125 (1996)
26.82 M. Grachten, J.L. Arcos, R.L. de Mántaras: Melody retrieval using the Implication/Realization model. In: *Proc. 6th Int. Conf. Music Inf. Retr* (Queen Mary University of London, London 2005)
26.83 R. Jackendoff: Musical parsing and musical affect, Music Percept. **9**(2), 199–229 (1991)
26.84 F. Lerdahl, C.L. Krumhansl: Modeling tonal tension, Music Percept. **24**, 329–366 (2007)
26.85 M. Hamanaka, K. Hirata, S. Tojo: Implementing a generative theory of tonal music, J. New Music Res. **35**, 249–277 (2006)
26.86 M. Hamanaka, K. Hirata, S. Tojo: FATTA: Full automatic time-span tree analyzer. In: *Proc. Int. Comput. Music Conf. (ICMC), Copenhagen* (2007) pp. 153–156
26.87 H. Riemann: *Musikalische Syntaxis. Grundriss einer harmonischen Satzbildungslehre* (Breitkopf Härtel, Leipzig 1877)
26.88 T. Winograd: Linguistics and computer analysis of tonal harmony, J. Music Theory **12**, 3–49 (1968)
26.89 J. Sundberg, B. Lindblom: Generative theories in language and music descriptions, Cognition **4**, 99–122 (1976)
26.90 M. Steedman: *The Syntactic Process* (MIT Press, Cambridge 2000)
26.91 M. Granroth-Wilding, M. Steedman: A robust parser-interpreter for jazz chord sequences, J. New Music Res. **43**, 355–374 (2014)
26.92 W.B. De Haas: *Music Information Retrieval Based on Tonal Harmony*, Doctoral Dissertation (Utrecht Univ., Utrecht 2012)
26.93 B. De Haas, M. Rohrmeier, R. Veltkamp, F. Wiering: Modeling harmonic similarity using a generative

grammar of tonal harmony. In: *Proc. 10th Int. Soc. Music Inf. Retr. Conf. (ISMIR 2009)*, ed. by M. Goto (2009) pp. 549–554

26.94 W.B. De Haas, J.P. Magalhães, F. Wiering: Improving audio chord transcription by exploiting harmonic and metric knowledge. In: *Proc. 13th Int. Soc. Music Inf. Retr. Conf. (ISMIR 2012)* (2012) pp. 295–300

26.95 D. Tidhar: *A Hierarchical and Deterministic Approach to Music Grammars and its Application to Unmeasured Preludes*, Ph.D. Thesis (Technische Universität Berlin, Berlin 2005)

26.96 E. Gilbert, D. Conklin: A probabilistic context-free grammar for melodic reduction. In: *Proc. Int. Workshop Artif. Intell. Music, 20th Int. Jt. Conf. Artif. Intell. (IJCAI), Hyderabad* (2007) pp. 83–94

26.97 R. Bod: Memory-based models of melodic analysis: Challenging the Gestalt principles, J. New Music Res. **30**, 27–37 (2001)

26.98 H. Schaffrath: The ESAC databases and MAPPET software, Comput. Musicol. **8**, 66 (1992)

26.99 H. Schaffrath: The ESAC electronic songbooks, Comput. Musicol. **9**, 78 (1994)

26.100 D. Marr: *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (W.H. Freeman, San Francisco 1982)

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

26.102 J.L. Elman: Finding structure in time, Cogn. Sci. **14**, 179–211 (1990)

26.103 R.N. Shepard: Structural representations of musical pitch. In: *Psychology of Music*, ed. by D. Deutsch (Academic Press, New York 1982) pp. 343–390

26.104 M. Rohrmeier, P. Rebuschat, I. Cross: Incidental and online learning of melodic structure, Conscious. Cogn. **20**(2), 214–222 (2011)

26.105 E. Servan-Schreiber, J.R. Anderson: Chunking as a mechanism of implicit learning, J. Exp. Psychol.: Learn. Mem. Cogn. **16**, 592–608 (1990)

26.106 M. Rohrmeier: *Implicit Learning of Musical Structure: Experimental and Computational Modelling Approaches*, Doctoral Dissertation (University of Cambridge, Cambridge 2010)

26.107 M. Rohrmeier, I. Cross: Tacit tonality: Implicit learning of context-free harmonic structure. In: *Proc. 7th Trienn. Conf. Eur. Soc. Cogn. Sci. Music*, ed. by J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, P.-S. Eerola (Univ. of Jyväskylä, Jyväskylä 2009) pp. 443–452

26.108 D.E. Rumelhart, G. Hinton, R. Williams: Learning internal representations through error propagation. In: *Parallel Distributed Processing: Experiments in the Microstructure of Cognition*, Vol. 1, ed. by D.E. Rumelhart, J.L. McClelland, PDP Research Group (MIT Press, Cambridge 1986) pp. 25–40

26.109 J.B. Pollack: Recursive distributed representations, Artif. Intell. **46**(1), 77–105 (1990)

26.110 E.W. Large, C. Palmer, J.B. Pollack: Reduced memory representations for music, Cogn. Sci. **19**(1), 53–96 (1995)

26.111 N. Boulanger-Lewandowski, Y. Bengio, P. Vincent: Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription. In: *Proc. 29th Int. Conf. Mach. Learn. (ICML), Edinburgh* (2012)

26.112 S. Cherla, T. Weyde, A. d'Avila Garcez, M.T. Pearce: Learning distributed representations for multiple-viewpoint melodic prediction. In: *Proc. 14th Int. Soc. Music Inf. Retriev. Conf. (ISMIR 2013), Curitiba* (2013)

26.113 M.T. Pearce, M. Rohrmeier: Music cognition and the cognitive sciences, TopiCS Cogn. Sci. **4**, 468–484 (2012)

26.114 A. Patel: *Music, Language and the Brain* (OUP, Oxford 2008)

26.115 A.D. Patel: Why would musical training benefit the neural encoding of speech? The OPERA hypothesis, Front. Psychol. **2**, 142 (2011)

26.116 S. Koelsch: *Brain and Music* (Wiley, Chichester 2012)

26.117 J.J. Bharucha, K. Stoeckig: Reaction time and musical expectancy: Priming of chords, J. Exp. Psychol. Hum. Percept. Perform. **12**(4), 403–410 (1986)

26.118 J.J. Bharucha, K. Stoeckig: Priming of chords: Spreading activation or overlapping frequency spectra?, Percept. Psychophys. **41**(6), 519–524 (1987)

26.119 B. Tillmann: Implicit investigations of tonal knowledge in nonmusicians listeners, Ann. N. Y. Acad. Sci. **1060**, 1–11 (2005)

26.120 E. Bigand, F. Madurell, B. Tillmann, M. Pineau: Effect of global structure and temporal organization on chord processing, J. Exp. Psychol. Hum. Percept. Perform. **25**(1), 184–197 (1999)

26.121 F. Marmel, F. Perrin, B. Tillmann: Tonal expectations influence early pitch processing: Evidence from ERPs, J. Cogn. Neurosci. **23**, 3095–3104 (2011)

26.122 E. Bigand, B. Tillmann, B. Poulin, D.A. D'Adamo: The effect of harmonic context on phoneme monitoring in vocal music, Cognition **81**, B11–B20 (2001)

26.123 E. Bigand, B. Poulin, B. Tillmann, D. D'Adamo: Cognitive versus sensory components in harmonic priming effects, J. Exp. Psychol. Hum. Percept. Perform. **29**(1), 159–171 (2003)

26.124 E. Bigand, B. Tillmann, B. Poulin-Charronnat, D. Manderlier: Repetition priming: Is music special?, Q. J. Exp. Psychol. Sect. A **58**(8), 1347–1375 (2005)

26.125 E. Bigand, B. Poulin, B. Tillmann, F. Madurell, D.A. D'Adamo: Sensory versus cognitive components in harmonic priming, J. Exp. Psychol. Hum. Percept. Perform. **29**(1), 159–171 (2003)

26.126 B. Tillmann, J.J. Bharucha, E. Bigand: Implicit learning of music: A self-organizing approach, Psychol. Rev. **107**, 885–913 (2000)

26.127 E.G. Schellenberg: Simplifying the implication-realisation model of melodic expectancy, Music Percept. **14**(3), 295–318 (1997)

26.128 W.F. Thompson, M. Stainton: Expectancy in Bohemian folk song melodies: Evaluation of implicative principles for implicative and closural intervals, Music Percept. **15**(3), 231–252 (1998)
26.129 P.T. von Hippel, D. Huron: Why do skips precede reversals? The effects of tessitura on melodic structure, Music Percept. **18**(1), 59–85 (2000)
26.130 F.A. Russo, L.L. Cuddy: A common origin for vocal accuracy and melodic expectancy: Vocal constraints, J. Acoust. Soc. Am. **105**, 1217 (1999)
26.131 A. Tierney, F.A. Russo, A.D. Patel: The motor origins of human and avian song structure, Proc. Natl. Acad. Sci. (PNAS) USA **108**, 15510–15515 (2011)
26.132 P.T. von Hippel: Melodic-expectation rules as learned heuristics. In: *Proc. 7th Int. Conf. Music Percept. Cogn.*, ed. by C. Stevens, D. Burnham, E. Schubert, J. Renwick (Causal Productions, Adelaide 2002) pp. 315–317
26.133 N. Oram, L.L. Cuddy: Responsiveness of Western adults to pitch-distributional information in melodic sequences, Psychol. Res. **57**(2), 103–118 (1995)
26.134 C.L. Krumhansl, P. Toivanen, T. Eerola, P. Toiviainen, T. Järvinen, J. Louhivuori: Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks, Cognition **76**(1), 13–58 (2000)
26.135 T. Eerola: Data-driven influences on melodic expectancy: Continuations in North Sami Yoiks rated by South African traditional healers. In: *Proc. 8th Int. Conf. Music Percept. Cogn*, ed. by S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, P. Webster (Causal Productions, Adelaide 2004) pp. 83–87
26.136 J.C. Carlsen: Some factors which influence melodic expectancy, Psychomusicology **1**(1), 12–29 (1981)
26.137 M.A. Castellano, J.J. Bharucha, C.L. Krumhansl: Tonal hierarchies in the music of North India, J. Exp. Psychol. Gen. **113**(3), 394–412 (1984)
26.138 E.J. Kessler, C. Hansen, R.N. Shepard: Tonal schemata in the perception of music in Bali and the West, Music Percept. **2**(2), 131–165 (1984)
26.139 C.L. Krumhansl, J. Louhivuori, P. Toiviainen, T. Järvinen, T. Eerola: Melodic expectation in Finnish spiritual hymns: Convergence of statistical, behavioural and computational approaches, Music Percept. **17**(2), 151–195 (1999)
26.140 M. Rohrmeier, R. Widdess: Incidental learning of melodic structure of North Indian music, Cogn. Sci. (2016), https://doi.org/10.1111/cogs.12404
26.141 E.G. Schellenberg, M. Adachi, K.T. Purdy, M.C. McKinnon: Expectancy in melody: Tests of children and adults, J. Exp. Psychol.: Gen. **131**(4), 511 (2002)
26.142 C.L. Krumhansl: *Cognitive Foundations of Musical Pitch* (Oxford Univ. Press, Oxford 1990)
26.143 W.E. Caplin: *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven* (Oxford Univ. Press, Oxford 1998)
26.144 D. Deutsch: The processing of structured and unstructured tonal sequences, Percept. Psychophys. **28**(5), 381–389 (1980)
26.145 M.G. Boltz, M.R. Jones: Does rule recursion make melodies easier to reproduce? If not, what does?, Cogn. Psychol. **18**(4), 389–431 (1986)
26.146 V.J. Konečni: Elusive effects of artists' "messages". In: *Cognitive Processes in the Perception of Art*, ed. by W.R. Crozier, A.J. Chapman (North Holland, Amsterdam 1984) pp. 71–93
26.147 H. Gotlieb, V.J. Konečni: The effects of instrumentation, playing style, and structure in the Goldberg Variations by Johann Sebastian Bach, Music Percept. **3**, 87–102 (1985)
26.148 N. Cook: The perception of large-scale tonal closure, Music Percept. **5**(2), 197–206 (1987)
26.149 R.O. Gjerdingen: An experimental music theory? In: *Rethinking Music*, ed. by M. Everist, N. Cook (Oxford Univ. Press, Oxford 1999) pp. 161–170
26.150 M. Karno, V.J. Konečni: The effects of structural interventions in the First Movement of Mozart's Symphony in G-Minor K. 550 on aesthetic preference, Music Percept. **10**, 63–72 (1992)
26.151 E.W. Marvin, A. Brinkman: The effect of modulation and formal manipulation on perception of tonic closure by expert listeners, Music Percept. **16**, 389–408 (1999)
26.152 S. Koelsch, M. Rohrmeier, R. Torrecuso, S. Jentschke: Processing of hierarchical syntactic structure in music, Proc. Natl. Acad. Sci. Am. **110**, 15443–15448 (2013)
26.153 S.M. Kosslyn: If neuroimaging is the answer, what is the question?, Proc. R. Soc. B **354**, 1283–1294 (1999)
26.154 M. Mather, J.T. Cacioppo, N. Kanwisher: How fMRI can inform cognitive theories, Perspect. Psychol. Sci. **8**, 108–113 (2013)
26.155 R.A. Poldrack: Can cognitive processes be inferred from neuroimaging data?, Trends Cogn. Sci. **10**, 59–63 (2006)
26.156 C.N. White, R.A. Poldrack: Using fMRI to constrain theories of cognition, Perspect. Psychol. Sci. **8**, 79–83 (2013)
26.157 S. Koelsch, T. Gunter, A.D. Friederici: Brain indices of music processing: "Nonmusicians" are musical, J. Cogn. Neurosci. **12**(3), 520–541 (2000)
26.158 S. Koelsch, S. Kilches, N. Steinbeis, S. Schelinski: Effects of unexpected chords and of performer's expression on brain responses and electrodermal activity, PLoS One **3**(7), e2631 (2008)
26.159 S. Koelsch, B.-H. Schmidt, J. Kansok: Effects of musical expertise on the early right anterior negativity: An event-related brain potential study, Psychophysiology **39**(5), 657–663 (2002)
26.160 S. Leino, E. Brattico, M. Tervaniemi, P. Vuust: Representation of harmony rules in the human brain: Further evidence from event-related potentials, Brain Res. **1142**, 169–177 (2007)
26.161 P. Loui, T. Grent, D. Torpey, M.G. Woldorff: Effects of attention on the neural processing of harmonic syntax in western music, Cogn. Brain Res. **25**, 678–687 (2005)
26.162 A.D. Patel, E. Gibson, J. Ratner, M. Besson, P.J. Holcomb: Processing syntactic relations in language and music: An event-related potential

study, J. Cogn. Neurosci. **10**, 717–733 (1998)
26.163 N. Steinbeis, S. Koelsch, J.A. Sloboda: The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological and neural responses, J. Cogn. Neurosci. **18**(8), 1380–1393 (2006)
26.164 S. Koelsch, E. Schroger, T.C. Gunter: Music matters: Preattentive musicality of the human brain, Psychophysiology **39**, 38–48 (2002)
26.165 S.G. Kim, J.S. Kim, C.K. Chung: The effect of conditional probability of chord progression on brain response: An MEG study, PLoS ONE **6**, e17337 (2011)
26.166 P. Loui, E.H. Wu, D.L. Wessel, R.T. Knight: A generalized mechanism for perception of pitch patterns, J. Neurosci. **29**(2), 454–459 (2009)
26.167 M. Besson, F. Faïta: An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians, J. Exp. Psychol. Hum. Percept. Perform. **21**, 1278–1296 (1995)
26.168 H. Nittono, T. Bito, M. Hayashi, S. Sakata, T. Hori: Event-related potentials elicited by wrong terminal notes: Effects of temporal disruption, Biol. Psychol. **52**, 1–16 (2000)
26.169 K.A. Paller, G. McCarthy, C.C. Wood: Event-related potentials elicited by deviant endings to melodies, Psychophysiology **29**(2), 202–206 (1992)
26.170 D. Schön, M. Besson: Visually induced auditory expectancy in music reading: A behavioral and electrophysiological study, J. Cogn. Neurosci. **17**, 694–705 (2005)
26.171 R. Verleger: P3-evoking wrong notes: Unexpected, awaited, or arousing?, Int. J. Neurosci. **55**(2–4), 171–179 (1990)
26.172 J.J. Bharucha: Music cognition and perceptual facilitation: A connectionist framework, Music Percept. **5**, 1–30 (1987)
26.173 R.A. Miranda, M.T. Ullman: Double dissociation between rules and memory in music: An event-related potential study, NeuroImage **38**(2), 331–345 (2007)
26.174 S. Koelsch, S. Jentschke: Differences in electric brain responses to melodies and chords, J. Cogn. Neurosci. **22**(10), 2251–2262 (2010)
26.175 S. Koelsch, T.C. Gunter, M. Wittfoth, D. Sammler: Interaction between syntax processing in language and in music: An ERP study, J. Cogn. Neurosci. **17**(10), 1565–1577 (2005)
26.176 E. Carrus, M.T. Pearce, J. Bhattacharya: Melodic pitch expectation interacts with neural responses to syntactic but not semantic violations, Cortex **49**, 2186–2200 (2012)
26.177 N. Steinbeis, S. Koelsch: Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns, Cereb. Cortex **18**(5), 1169–1178 (2008)
26.178 A.D. Friederici: Towards a neural basis of auditory sentence processing, Trends Cogn. Sci. **6**, 78–84 (2002)
26.179 E. Fonteneau, H.K.J. van der Lely: Electrical brain responses in language-impaired children reveal grammar-specific deficits, PLoS One **3**(3), e1832 (2008)
26.180 S. Jentschke, S. Koelsch, S. Sallat, A.D. Friederici: Children with specific language impairment also show impairment of music-syntactic processing, J. Cogn. Neurosci. **20**(11), 1940–1951 (2008)
26.181 A.D. Patel, J.R. Iversen, M. Wassenaar, P. Hagoort: Musical syntax processing in agrammatic Broca's aphasia, Aphasiology **22**, 776–789 (2008)
26.182 B. Maess, S. Koelsch, T.C. Gunter, A.D. Friederici: 'Musical syntax' is processed in Broca's area: An MEG-study, Nat. Neurosci. **4**, 540–545 (2001)
26.183 S. Koelsch, T.C. Gunter, D.Y. von Cramon, S. Zysset, G. Lohmann, A.D. Friederici: Bach speaks: A cortical 'language-network' serves the processing of music, Neuroimage **17**, 956–966 (2002)
26.184 B. Tillmann, P. Janata, J.J. Bharucha: Activation of the inferior frontal cortex in musical priming, Cogn. Brain Res. **16**, 145–161 (2003)
26.185 D. Levitin, V. Menon: Musical structure is processed in "language" areas of the brain: A possible role for Brodmann area 47 in temporal coherence, NeuroImage **20**, 2142–2152 (2003)
26.186 S. Brown, M.J. Martinez, L.M. Parsons: Music and language side by side in the brain: A PET study of the generation of melodies and sentences, Eur. J. Neurosci. **23**, 2791–2803 (2006)
26.187 T.R. Knösche, C. Neuhaus, J. Haueisen, K. Alter, B. Maess, A.D. Friederici, O. Witte: The perception of phrase structure in music, Hum. Brain Mapp. **24**, 259–273 (2005)
26.188 K. Steinhauer, K. Alter, A.D. Friederici: Brain potentials indicate immediate use of prosodic cues in natural speech processing, Nat. Neurosci. **2**, 191–196 (1999)
26.189 K. Steinhauer, A.D. Friederici: Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers, J. Psycholinguist. Res. **30**, 267–295 (2001)
26.190 A. Pannekamp, U. Toepel, K. Alter, A. Hahne, A.D. Friederici: Prosody-driven sentence processing: An event-related brain potential study, J. Cogn. Neurosci. **17**, 407–421 (2005)
26.191 C. Neuhaus, T. Knösche, A. Friederici: Effects of musical expertise and boundary markers on phrase perception in music, J. Cogn. Neurosci. **18**, 472–493 (2006)
26.192 P.N. Johnson-Laird: Jazz improvisation: A theory at the computational level. In: *Representing Musical Structure*, ed. by P. Howell, R. West, I. Cross (Academic Press, London 1991) pp. 291–325
26.193 C. Francois, D. Schön: Musical expertise boosts implicit learning of both musical and linguistic structures, Cereb. Cortex **21**, 2357–2365 (2011)
26.194 I. Peretz, E. Brattico, M. Järvenpää, M. Tervaniemi: The amusic brain: In tune, out of key, and unaware, Brain **132**, 1277–1286 (2009)
26.195 M. Rohrmeier, S. Koelsch: Predictive information processing in music cognition. A critical review, Int. J. Psychophysiol. **83**(2), 164–175 (2012)
26.196 S. Koelsch: Towards a neural basis of processing musical semantics, Phys. Life Rev. **8**(2), 89–105 (2011)

26.197 U. Reich: The meanings of semantics: Comment on "Towards a neural basis of processing musical semantics" by Stefan Koelsch, Phys. Life Rev. **8**(2), 120–121 (2011)
26.198 L.B. Meyer: *Emotion and Meaning in Music* (University of Chicago Press, Chicago 1956)
26.199 E. Hanslick: *Vom Musikalisch-Schönen* (R. Weigel, Leipzig 1854), Reprint: Darmstadt 1965
26.200 L.B. Meyer: Meaning in music and information theory, J. Aesthet. Art Crit. **15**(4), 412–424 (1957)
26.201 I.H. Witten, L.C. Manzara, D. Conklin: Comparing human and computational models of music prediction, Comput. Music J. **18**(1), 70–80 (1994)
26.202 P.N. Juslin, D. Västfjäll: Emotional responses to music: The need to consider underlying mechanisms, Behav. Brain Sci. **31**(5), 559–575 (2008), discussion pp. 575–621