25. Musical Syntax I: Theoretical Perspectives

Martin Rohrmeier, Marcus Pearce

The understanding of musical syntax is a topic of fundamental importance for systematic musicology and lies at the core intersection of music theory and analysis, music psychology, and computational modeling. This chapter discusses the notion of musical syntax and its potential foundations based on notions such as sequence grammaticality, expressive unboundedness, generative capacity, sequence compression and stability. Subsequently, it discusses problems concerning the choice of musical building blocks to be modeled as well as the underlying principles of sequential structure building. The remainder of the chapter reviews the main theoretical proposals that can be characterized under different mechanisms of structure building, in particular approaches using finite-context or finite-state models as well as tree-based models of context-free complexity (including the Generative Theory of Tonal Music) and beyond. The chapter concludes with a discussion of the main issues and questions driving current research and a preparation for the subsequent empirical chapter *Musical Syntax II*.

> *The idea that there is a grammar of music is probably as old as the idea of a grammar itself. Mark Steedman* [25.1, p. 1]

25.1 Outline

What distinguishes music from other sounds? One answer is to be found in the manner in which the elements are organized and related within a structural framework and, most importantly, the apprehension of this structure by a listener, so that the sound is *experienced as music* by that listener. Therefore, discovering the principles of musical structure building is one of the central questions for theoretical and empirical music research. Despite the strong historical (and methodological) divide between music-theoretical, computational, and psychological/neuroscientific approaches, questions about musical structure and the perception of it facilitate a close link across traditional divisions between disciplines [25.2]. Note that we use the term *computational* to describe a theory that is expressed in computational terms, whether or not it is actually implemented as a computer program.

Exploring the principles of musical structure building naturally requires us to distinguish between the goal of uncovering rules governing the structure of music (an

R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, https://doi.org/10.1007/978-3-662-55004-5_25

external goal) and the cognitive principles of the perception and production of these structures (an internal psychological goal). Yet both aspects form two sides of the same coin: the capacities and limitations of human perception and cognitive processes influence the possible structures that composers can use (for a similar argument about language, the reader is directed to [25.3]) and, together with other constraints (e.g., those imposed by cultural factors or the physical properties of instruments, constraints of the hands or the body, constraints of the performance and so on [25.4]) give rise to the musical structures that we find in music. In turn, musical structure is acquired implicitly by listeners from mere exposure and musical interaction and represented internally [25.5, 6] and, ultimately, reproduced in compositional practice (since composers are listeners before they become composers). However, there can be no learning without a hypothesis space and therefore theoretical models of musical structure, especially those grounded in computational modeling, provide a useful approach to understanding the hypothesis space that human learners are faced with when they acquire the syntactic structure of a musical style.

Finding a formal characterization of musical structure brings traditional music theory in close connection with computational modeling since the search for an optimal structural description (that relates to structures as *heard*) strongly implies *modeling* the internal structure of the music (whether it is a single composition, a part thereof, or a corpus). Since music is an inherently psychological phenomenon, we often use psychological understanding to guide the development of structural models of music, just as we use structural models of music to guide the development and testing of psychological theories of the perception of musical structure.

The disciplines involved in research on musical syntax range from musicology and music theory, through computational modeling, to psychology and neurobiology. Although the disciplinary perspectives are distinct (e.g., it is possible to develop a structural theory that is optimal according to some criterion, such as simplicity, but not according to the criterion of matching human perception and cognition), in this contribution, we focus on a converging picture that emerges when musical syntax is examined by triangulating between theory, computational modeling, and cognitive research. Here we focus on theoretical approaches to musical syntax while empirical research using computational models, psychological experimentation, and neuroimaging are covered in the companion chapter, *Musical Syntax II.*

25.2 Theories of Musical Syntax

25.2.1 The Concept of Musical Syntax

Berwick et al. [25.7, p. 89] give a brief account of syntax as:

> *the rules for arranging items (sounds, words, word parts, phrases) into their possible permissible combinations in a language.*

In human language the set of items (alphabet of symbols) may be words and morphosyntactic units, in birdsong they may be pitches, slides and other sounds. In music the symbols may be melodic notes, chords, voice-leading patterns or relationships between voices, timbral qualities and so on. Many music-theoretical approaches constitute informal, verbal accounts of syntactic models of music. Although the use of strict and well-defined formalisms is not (yet) common in music theory, there are some accounts that employ the notion of syntax in music theory. For instance, *Aldwell* and *Schachter* write the following in order to characterize *harmonic syntax* [25.8, p. 139]:

> *One way that music resembles language is that the order of things is crucial in both.* I went to the concert *is an English sentence, whereas* I concert went the to *is not. Similarly, I-VII6-I6-II6-V7-I* [...] *is a coherent progression of chords, whereas I-I6-VII6-II6-I-V7* [...] *is not, as you can hear if you play through the two examples. In the study of language, the word* syntax *is used to refer to the arrangement of words to form sentences; word order is a very important component of syntax. In studying music, we can use the term* harmonic syntax *to refer to the arrangement of chords to form progressions; the order of chords within these progressions is at least as important as the order of words in language. (Other components of harmonic syntax are the position of chords within phrases, the preparation and resolution of dissonances, and the relation of chord progressions to melody and bass lines.)*

A syntactic theory might be applied to any aspect of musical structure – melody, harmony, rhythm, metre, grouping structure, form, or even aspects such as timbre and dynamics. In practice, syntactic approaches have typically been applied to *what* happens in a musical sequence – e.g., predicting (combinations of) pitches and chords – rather than *when* it happens. Conversely, theories of rhythm and metre often do not take an explicitly syntactic approach. By analogy, metrical and rhythmic features of language are often studied from the perspective of phonology rather than syntax. A well-formed harmonic sequence, for instance, may be assigned to metrical structure in a regular or irregular way. It is important to note that despite the predominance of Western music in theoretical and cognitive research [25.9], the general notion of musical syntax is *not* limited to Western tonal music – and there are approaches addressing non-Western music [25.10–12]. Different aspects of musical structure may be more or less important in different musical styles and cultures.

Several models for musical syntax have been proposed based on different levels of structural representation (melodic structure, harmony and chords, bass lines, outer voices, voice-leading, other types of categorized sound, polyphonic pitch structure and so on). Here, we reserve the term *syntax* for approaches presenting a formal system characterizing the sequential structure of such building blocks, in contrast to the more general term *musical structure* which captures the rich interaction of different musical features such as rhythm, metre, timbre, counterpoint, dynamics, phrasing, instrumentation, agogics, and so on. The precise identity of these building blocks is one of the central ongoing research questions in musical syntax.

The general term *musical structure* refers to the way in which one or more pieces of music may be represented in terms of their constituent parts, potentially reflecting a wide range of different musical features including rhythm, metre, timbre, counterpoint, dynamics, phrasing, instrumentation, agogics, and so on.

Musical syntax is a formal characterization of the principles governing permissible sequential structure in music. It characterizes sequences of musical events generated from a *lexicon of building blocks* and a set of *rules* governing how the building blocks are combined.

The *lexicon* (the set of building blocks) may consist of single events or patterns (schemata) of notes, glissandi, rests, chords, voice-leading patterns, timbres, or other noises. The *rules* may constitute any formal system that characterizes how sequences may (and may not) be formed by combining elements from the lexicon.

25.2.2 Foundations of Musical Syntax

Why do we need a syntax of music? When characterizing musical structure, and the cognitive representation and processing of that structure, several issues arise which motivate the development of a formal syntactical understanding of music. These include distinguishing regular and irregular musical structures (i. e., making grammaticality judgements), the fact that the space of possible musical compositions is theoretically infinite (or unbounded), the idea that we want to be able to describe structural relationships within musical sequences (i. e., focus on strong generative capacity compared to weak generative capacity). Syntax is also relevant to tasks such as compression, identifying the stability of events at different levels within music, and measuring musical similarity. We investigate these issues further in the following sections.

Grammaticality

One core foundation for the concept of musical syntax is the notion of regularity, permissibility, well-formedness or grammaticality, i. e., the characterization of structures that are regular or irregular with respect to a particular system (representing, for example, a musical style). If such a distinction were irrelevant, the characterization of musical syntax would be unnecessary since every sequence would be equally plausible. However, musical styles and idioms are implicitly characterized by regular and irregular sequences. Although categorical grammaticality decisions are often made, the distinction may be one of degree (compare [25.13–15], in linguistics). For instance, not every chord sequence or every musical form is regular in the 18th century Classical idiom [25.16]. Another example illustrates a regular and irregular common-practice harmonic sequence (Fig. 25.1 adapted from [25.17]). Musical regularity can be characterized empirically through (computational or hand-conducted) corpus analysis, which can provide information about frequent and less frequent regular patterns and indirectly about irregularity by the discovery of absent and low-probability patterns (although absence does not necessarily constitute evidence for ungrammaticality). Grammaticality can also be experimentally established through psychological experiments. Furthermore, introspective analyses by individual experts may be regarded as single-participant experiments, with some extrapolation to wider groups (whether expert or otherwise) assumed. In this context it is essential to understand the importance of negative evidence in the form of explicit instruction about implausible or irregular structures (note that this is different from the absence of positive evidence). While some regularities and rules may be inferred from positive data alone (i. e., the presence of well-formed structures), it is negative evidence that makes the strongest conclusions possible with regard to range, scope of generalization, and mutual interaction of grammatical systems. There is continuing debate about whether and how people might receive negative evidence in the development of language, but this topic has received little attention in the domain of music.

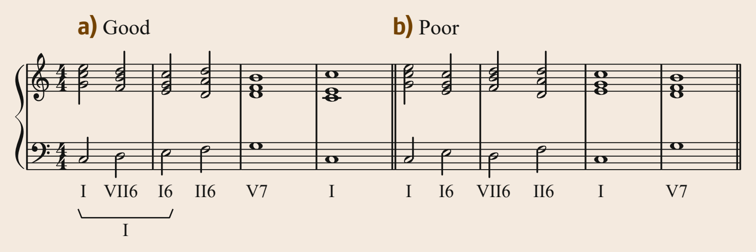

Fig. 25.1a,b The contrast between a good **(a)** and a poor **(b)** harmonic progression as discussed by *Aldwell* and *Schachter* [25.8, p. 140]. **(b)** is poor because the dependencies between the chords are disrupted – for instance, the II6 chord is not functionally well connected to its context (even though it features good voice-leading). Further note that in the analysis of the good example, the authors propose a hierarchical analysis of I VII6 I6 as a prolongation of a single I harmony

Unboundedness

The set of possible musical structures is unbounded – music, in Humboldt's famous words, makes infinite use of finite means. It is simple to demonstrate the unboundedness of musical structures: for every sequence we can imagine a longer one or a variation of it that inserts another element (tone, chord, etc.) into the sequence; we can further imagine a composition that never ends (such as ideal airport music). Hence, it is impossible to construct an exhaustive list of all musical sequences. Therefore, the only way to characterize musical structure is by employing a finite set of building blocks and recursive (or iterative) rules to generate grammatical sequences based on the recombination of building blocks using the rules. Generative grammars [25.18, 19] are *one kind* of formalism that embodies this principle, often used in theoretical approaches to the syntax of music (and other auditory sequences such as language or birdsong). Note that in this context the term *generative* does not refer to (human) music generation but to the description and analysis of a set of sequences by formal rules that are capable of generating them by a well-defined formal mechanism (such as a formal grammar).

Weak and Strong Generative Capacity

A syntactic model of a set of (musical) sequences may focus on the description of the surface sequences in order to reproduce exactly those sequences. For a given language (i. e., a set of strings), such a characterization may be accurate in terms of coverage (i. e., they can generate the set of strings). This is referred to as *weak generative capacity.* However, for most languages there is an infinite number of formal models that adequately describe the language, many of which are highly implausible. For this reason it is desirable that a syntactic theory matches theoretical insights as well as cognitively relevant (or adequate) structures, provides useful and testable generalizations, and achieves optimal compression (see below). This broader concept is known as *strong generative capacity*.

Compression

Characterizing a set of (musical) sequences using a generative theory allows us to capture a potentially infinite set of sequences using a finite set of rules. In this sense, we can think of generative theories in terms of the extent to which they enable compression through efficient representations of a set of sequences. Highly efficient, sparse encoding of the environment constitutes a core principle of cognitive systems [25.20, 21], and there is a close relationship between prediction and compression because we only need to store the information that is not predictable using a model [25.22]. Research in music information retrieval [25.23] and music psychology [25.24] has used general-purpose compression algorithms as models of musical complexity. A model that better captures structural regularities with general coverage in a given musical idiom is expected to be capable of more accurate prediction and, therefore, compression. Conversely, we can use compressibility (of unseen data to ensure generalizability) as a measure of the power and efficiency of a generative theory (and the latent structure that it postulates). However, a more complex model will itself consume more space, meaning that increased level of compression of the data must exceed the increased size of the model in order to be efficient. In this respect, approaches following the paradigms of minimum description length (MDL; [25.25]) and Bayesian model comparison provide closely related methods [25.22, Chap. 28] for comparing different candidate models taking into account differences in their complexity and the numbers of free parameters and so on [25.26, 27]. See *Mavromatis* [25.28] for an example of these principles applied to music.

Compression-based evaluation of a model of musical syntax is independent from other questions such as grammaticality or weak/strong generative capacity. In particular, the criterion of optimal compression makes it possible to evaluate and compare syntactic models independently of the grammaticality distinction as well as independently of tests (such as pumping lemmata) that require grammaticality distinctions over sequences that are extremely improbable and do not generalize over corpora (such as n-th level center embeddings). In this context, compression provides a better way to provide a foundation for strong generative capacity and also assess the cognitive relevance of a proposed syntactic account of a (musical) language.

Stability, Similarity, and Semantics as Underpinnings of Syntax

There are several other ways to motivate syntactic structure in music. One of these is the proposal that we need an account of syntactic structure in music to be able to predict the relative stability of musical events. Many music theorists observe that in harmonic, melodic or voice-leading sequences, some events may be considered ornamental or accidental whereas others are structurally fundamental [25.8, 29, 30]. If this notion of relative structural stability – not to be confused with tonal stability and the tonal hierarchy [25.31] – is extended to a fully recursive structure (i. e., not just to individual notes or chords but also to motifs, phrases, and other larger scale components of musical form), it can be accounted for using a hierarchical syntactic formalism. Whether or not this type of structure is in turn coextensive with the above forms of establishing hierarchical structure remains a question open for further theoretical investigation.

Another, related avenue for establishing hierarchical structure is similarity. From a theoretical and psychological perspective musical similarity may be construed in terms of operations of omission or insertion of events with respect to a common core structure (for instance, differences between different cover versions of a song). In this context, it is important that such operations respect (hierarchical) structural boundaries of constituents (e.g., a tonic expansion) rather than comparisons between unstructured surface sequences. For example, *De Haas* et al. [25.32] implemented a similarity measure that is closely related to structural stability in terms of the largest common embeddable subtree between two compositions. This approach outperformed *edit distance* (a structure-free surface comparison between sequences) in predicting harmonic similarity between music sharing similar melodies. Similarity is also closely related to the concept of compression since we can train a syntactic model on one piece of music and use that model to predict another piece of music – greater degrees of predictability (and hence compressibility) indicate greater degrees of structural overlap between the pieces [25.21, 23].

Finally, semantics may constrain syntactic structures, particularly in linguistics. Whereas linguistic syntactic structures to a large extent serve the temporalization/linearization of semantic structure (in terms of form/meaning pairs), there is no immediate analogy in music. Although music may express meaning in terms of illocutionary acts like warnings, or aggression, or in terms of symbolic associations, it is agreed that music, in general, lacks complex, explicit propositional semantic forms ([25.33] and its discussion [25.33–36]). However, the patterns of relative stability outlined above (which are themselves related to syntactic structure) lead to perception and experience of tension and release by the listener, which can be viewed as a kind of semantic interpretation [25.37–40]. However, further research is required to examine these potential relationships between syntax and semantics in music.

25.3 Models of Musical Syntax

A model of musical syntax consists of two core components: first, a choice of the underlying representation for musical building blocks and how they relate to the musical surface; second, a formalism used to generate musical structure based on the set of building blocks.

25.3.1 Building Blocks

The choice of building block is fundamental for the syntactic model. In contrast to language, where the set of morphosyntactic features is largely accepted, syntactic models of music have made different choices of building block. This entails modeling musical structure at different levels of representation (or abstraction), such as: harmony and chord sequence, bass line, melodic line (diastematics), outer voices and voice-leading, or polyphonic pitch structure. Every choice involves selecting a distinction between structural and nonstructural items *with respect to the underlying model*. For instance, a model of harmonic syntax may regard different surface and melodic realizations of a chord sequence as equivalent; similarly, a theory of voice-leading would regard certain note repetitions, trills, or ornaments as nonstructural. Given the divergence of representations, styles, and level of abstraction adopted by different approaches in the literature, there is no consensus at present how (and based on which principles) a *fundamental domain* of building blocks for a musical syntax could be established independently of the modeling goals.

We should mention here that some very interesting work has been done on representational spaces for various aspects of musical elements including, most notably, pitch spaces [25.37, 41–48] and metrical structure [25.49]. These theories define how these aspects of music may be expressed in algebraic ways and potentially represented by cognitive systems [25.50], but since they characterize the formal space of musical objects rather than explicitly specifying how sequences of elements may be combined, we do not consider them as theories of syntax proper.

25.3.2 Structure Building

Traditionally there have been a number of theoretical attempts to characterize the sequential structure of elements in a sequence, ranging from Markov models to context-free languages and corresponding probabilistic models. Many theories of structure have used explicit types of formal languages in the Chomsky hierarchy and its extensions [25.51]. Characterizations with models of different complexity involve a trade-off between the expressive (and compressive) power of the representation and corresponding processing requirements.

The languages generated by each class of grammar form proper subsets of the languages generated by classes of grammar higher up in the hierarchy. However, as we move up the hierarchy, the complexity of recognition and parsing increases in tandem with the increased expressive power of each class of grammar. In particular, while context-free grammars (and those higher in the hierarchy) are capable of capturing phenomena, such as embedded structure, which cannot be captured by finite-state grammars, they also bring with them many problems of intractability and undecidability, especially in the context of grammar induction [25.52].

It is fundamental to note that the Chomsky hierarchy and its extensions [25.51] constitute just *one way* of characterizing (musical) sequential structure. They are in no particular way primary or more natural than other approaches that characterize classes of infinite sets of strings, except in historical terms. There are many ways to characterize sequential structure, as any handbook of formal languages demonstrates (e.g., Handbook formal languages, [25.53]). Furthermore, computational models are, fundamentally, in no respect distinct from hand-crafted models by theorists in terms of their expressive power [25.5, 54, 55].

25.4 Syntactic Models of Different Complexity

25.4.1 Finite-Context Models

There is an interesting subclass of grammar contained within the class of finite-state grammars which are known as *finite context* grammars [25.56, 57]. In finite context automata, the next state is completely determined by testing a finite portion of length $n-1$ of the end of the already processed portion of the input sequence [25.57]. The core idea of these very local models is to characterize sequential structure by identifying possible element-to-element transitions (how elements may follow or precede each other). This characterization formally amounts to a table that lists grammatical relationships between each possible combination of elements (such as chord, note, or root transitions). This account is easily extended to larger context-lengths: the next element may be related not only to its predecessor, but also to the sequence of 2, 3, or more preceding elements. What such accounts have in common is the assumption that there are no (unbounded) dependencies between events longer than the relevant context of the model. In general, finite-context models correspond to the formal subcategory of strictly local languages (*k*-factor languages) and are also referred to as Markov or *n*-gram models.

A *k-factor language* is formally defined by a set of factors (strings of length k). A sequence is grammatical iff every subsequence of length k is part of the set of factors. Several models have been proposed in music theory and cognition that contain, in part, *k-factor* models [25.58–60]. It is important to note here that schema-theoretic approaches [25.61–63] do *not* naturally correspond with *k-factor languages* without modification (since they involve reductions, nonlocal patterns, and the ability to distinguish notes that are structurally important from those that are not).

These characterizations of structure, however, only draw a distinction between regular and irregular sequences, yet within those categories, they consider every possible sequence equally. For many theoretical purposes this is insufficient as some of these structures occur very frequently whereas other transitions are rare, less common, or unlikely. This theoretical requirement demands a characterization that is based not only on grammaticality but also on probability. It is straightforward to expand the above definition to incorporate probabilities: every entry in a transition matrix is associated with a probability. Probabilistic instantiations of the approach, therefore form a superset of the nonprobabilistic versions, which only allow the probabilities 0 (nongrammatical) and 1 (grammatical). Often these probabilities are estimated through analysis of frequency counts of events in a corpus [25.64, 65]. Such probabilistic extensions of k-factor languages are referred to as Markov models or n-gram models (for which n-grams correspond to probabilistic versions of k-factors). In an n-gram model, the sequence $e^{j}_{(j-n)+1}$ is called an n-gram (note that the subscript and superscript symbols denote the beginning and ending of a subsequence in the string; in the previous case it refers to the subsequence, from index $(j-n)+1$ to the index j) which consists, conceptually, of an initial subsequence, $e^{j-1}_{(j-n)+1}$, of length $n-1$ known as the *context* and a single symbol extension, e_j, called the *prediction*. The quantity $n-1$ is the *order* of the n-gram rewrite rule.

Such models are frequently used in computational models of music (see below), and also some music theoretical accounts (e.g., Piston's table of common root progressions, shown in Table 25.1).

By definition all types of strictly local or Markov models share the Markov assumption (25.1) and (25.2): the grammaticality (gr) of a subsequence or the probability (p) of a symbol appearing in a sequence depends only on its immediate preceding context of length k. This assumption means that these models cannot represent any nonconsecutive dependencies between musical elements beyond a fixed finite length.

$$gr\left(e^{i}_{1}\right) = gr\left(e^{i}_{i-n+1}\right) \tag{25.1}$$

$$p\left(e^{i}_{1}\right) \approx p\left(e^{i}_{i-n+1}\right) \tag{25.2}$$

Markov models provide powerful approximations to sequential structure for numerous practical applications independently of whether those sequences obey the Markov assumption. Nonetheless, such models are theoretically as well as practically limited in the extent to which they can capture and represent more complex structural features such as nonlocal dependencies, nested structures, and cross-serial dependencies. To some extent these limitations can be addressed by using sophisticated representation schemes such as the multiple-viewpoint formalism [25.64, 65] that extends the range of context that a Markov model can take into account by combining several Markov models over different feature spaces and (possibly) time scales, including nonadjacent events.

Table 25.1 Table of common root chord progressions (after [25.60])

	Is often followed by	Sometimes by	Less often by
I	IV or V	VI	II or III
II	V	IV or VI	I or III
III	VI	IV	I, II or V
IV	V	I or II	III or VI
V	I	VI or IV	III or II
VI	II or V	III or IV	I
VII	III	I	

25.4.2 Finite-State Models

Several theoretical approaches can be viewed as having equivalent representational power to *finite-state or regular grammars* in Chomsky's terminology. In contrast to *k-factor languages*, such models involve grammars that distinguish between (hidden) variables (nonterminal symbols) and surface symbols (terminals). Accordingly, *regular grammars* (i. e., grammars that only have rules of the form $A \rightarrow aB$; in which a refers to a terminal and A, B to nonterminals; see the appendix below) characterize sequential structure by building up a string from left to right. They form a true superset of *k-factor* languages. The formal machine that recognizes the set of strings generated by such a grammar is a *finite-state automaton* (informally, a flow-chart). The probabilistic counterpart to a regular grammar is the *Hidden Markov Model* (HMM; [25.66]).

25.4.3 Context-Free or Equivalent Models

There are several accounts of structure in music theory which go beyond the expressive power of finite-context and finite-state grammars (for further discussion [25.38]):

- Differences of structural importance
- Dependency structure, preparation, and ornamentation
- Headedness
- Nested structures
- Functional categories.

A useful starting point is the insight that the elements in a sequence may differ in structural importance, i. e., some can be left out without impairing grammaticality whereas others cannot. An early account by *Kostka* and *Payne* [25.67] refers to this as *levels* of harmony (note, however, that the observation is not restricted to harmony). Second, musical structure expresses dependencies: e.g., in a I II V or I III IV progression, the II or III chord may be understood as preparation for V or IV and not simply a sequential succession of I; accordingly, it is *dependent* on V or IV, not on I. This is expressed by the rules V → II V or *IV* → III IV (for further details the reader is directed to [25.68]). This further entails that goals are structurally more fundamental than their preparations and, conversely, that ornamentation and variation adds new material to basic structure. This notion of dependency structure further entails a notion of *headedness*, namely that in the II V progression, V is the fundamental chord, i. e., the head (as expressed in the left-hand side of the rules above).

Another central formal concept concerns nested structures: the notion of dependency introduced above may lead to the formation of dependent subsubsequences within a dependent subsequence within a sequence (and so on). For instance, the II chord (which is the preparation of the V chord) in the above sequence may be further elaborated, ornamented, or prepared. This leads to recursive structures in the form of tail recursion (chains) and nested recursion (sequence in a sequence). One prominent example of nested structure in tonal music is modulation (e.g., an early account by [25.69]; the reader is also directed to [25.38, 68, 70]; for nested structure in music see [25.1, 38, 71, 72]). Finally, *Riemann* [25.73] introduced the notion that chords can be classified into different functional categories (such as tonic, dominants, and subdominants) that may be functionally interchangeable, such as II and IV leading to V. Riemann considered harmonic sequences to be hierarchical [25.74], and *Rohrmeier* [25.38, 68] developed a context-free formalism for the functional approach to harmony. In this formalism, tree structures represent different harmonic sequences that fulfill identical harmonic functions in the same way in higher parts of the tree.

Context-free languages and hierarchical tree-based accounts are well-suited for representing these kinds of structural dependencies in sequences. A number of theories account for music in such theoretical terms: *Schenker* [25.75], *Lerdahl* and *Jackendoff* [25.71]; *Keiler* [25.77, 78]; *Steedman* [25.1, 72]; *Narmour* [25.30]; *Lerdahl* [25.37]; *Tidhar* [25.79]; *Rohrmeier* [25.38, 68]. Various partial or full computational implementations of these theories exist, as discussed below.

Schenker [25.75] proposed a theoretical account of music based on reductional analysis that reveals different layers of musical structure ranging from surface to foreground, middle ground and Ursatz. Briefly construed, his account entails that principles of counterpoint (such as neighbor notes) may be used to distinguish the structural importance of musical events.

Lerdahl and *Jackendoff*'s Generative Theory of Tonal Music [25.71] (GTTM) provides an account that brings the ideas expressed by Schenker into a rule-based theoretical framework, inspired by the generative approach to grammar in linguistics. It is, for example, founded on the assumption that a piece of music can be partitioned into hierarchically organized segments which may be derived through the recursive application of the same rules at different levels of the hierarchy. Specifically, the theory is intended to yield a hierarchical, structural description of the final cognitive state of an experienced listener to that composition.

According to GTTM, a listener unconsciously infers four types of hierarchical structure in a musical surface (Fig. 25.2): first, *grouping structure* which corresponds to the segmentation of the musical surface into units (e.g., motives, phrases, and sections); second, *metrical structure* which corresponds to the pattern of periodically recurring strong and weak beats; third, *time-span reduction* which represents the relative structural importance of pitch events within contextually established rhythmic units; and finally, *prolongational reduction* reflecting patterns of tension and relaxation amongst pitch events at various levels of structure. According to the theory, grouping and metrical structure are largely derived directly from the musical surface and these structures are used in generating a time-span reduction which is, in turn, used in generating a prolongational reduction. Each of the four domains of organization is subject to *well-formedness rules* that specify which hierarchical structures are permissible and which themselves may be modified in limited ways by *transformational rules*. While these rules are abstract in that they define only formal possibilities, *preference rules* select which well-formed or transformed structures ac-

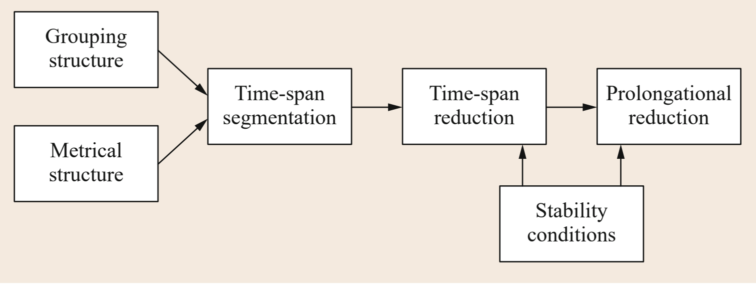

Fig. 25.2 The overall structure of GTTM (after [25.76, Fig. 10.6])

tually apply to particular aspects of the musical surface. Time-span and prolongational reduction additionally depend on tonal-harmonic *stability conditions* which are internal schemata induced from previously heard musical surfaces.

When individual preference rules reinforce one another, the analysis is stable and the passage is regarded as stereotypical whilst conflicting preference rules lead to an unstable analysis causing the passage to be perceived as ambiguous and vague. In this way, according to GTTM, the listener unconsciously attempts to arrive at the most stable overall structural description of the musical surface. Experimental studies of human listeners have found support for some of the preliminary components of the theory including the grouping structure [25.80] and the metrical structure [25.31].

The theory constitutes a formal predecessor to *Jackendoff*'s later parallel architecture framework of language [25.81, 82]. It is important to observe that the GTTM is not a *grammar* or a *syntax* of music: it provides a model of parsing but contains no generative rules or mechanisms to derive the musical surface, further it does not model a distinction between regular and irregular sequences. Rather than generating the musical surface, the GTTM is a theory of musical processing with only limited applicability as a theory of structural syntax per se.

It is highly challenging to develop formal context-free grammars that account for musical surface structure but several efforts have been made (e.g., [25.83, 84] for reviews). *Johnson-Laird* [25.85] used grammatical formalisms to investigate what has to be computed to produce acceptable rhythmic structure, chord progressions, and melodies in jazz improvisation. While a finite-state grammar (or equivalent procedure) can adequately compute the melodic contour, onset, and duration of the next note in a set of Charlie Parker improvisations, its pitch is determined by harmonic constraints derived from a context-free grammar modeling harmonic progressions. In a more recent approach, *Rohrmeier* [25.38, 68] introduces a set of context-free rules modeling the main features of tonal harmony from the common-practice period.

Context-free languages (and more complex formalisms) constitute supersets of regular and supraregular languages. In fact, the latter constitute local boundaries of context-free languages (i. e., substrings that do not use the features of nested embedding are regular; it is further possible to derive precise models of local transitions from context-free models). Accordingly, the distinction between these types of languages does not imply a forced alternative – rather, context-free language models can result from the addition of the above structural features to regular language accounts. Put another way, we can add degrees of context-free character to regular grammars.

25.4.4 Beyond Context-Free Complexity

Are there aspects of musical structure that require greater than context-free power to be modeled? Debates in theoretical linguistics of the past 25 years have reached a fairly consensical view that human language is mildly context sensitive [25.86, 87]. It requires syntactic power that is stronger than context-free but considerably less strong than the immense computational power of full context-sensitive grammars. One example of this context-sensitive complexity is given by cross-serial dependencies (as in Dutch or Swiss German relative clauses [25.86, 88]) that cannot be expressed by context-free grammars. In the Chomskian tradition, minimalist grammars, that are equally mildly context-sensitive [25.89], adopted two mechanisms of external merge (similar to a context-free tree building operation) and internal merge (combining an already derived branch of a tree with different free positions in the tree). Internal merge may express features such as movement (*Sue wondered which book Peter read?*). *Katz* and *Pesetsky* [25.90] argue that musical and linguistic structure are formally equivalent in the sense that both require structure-building operations based on external and internal merge.

What about music? In his review, *Roads* [25.83] argues that the strict hierarchy characteristic of context-free grammars is difficult to reconcile with the ambiguity inherent in music. Faced with the need to consider multiple attributes occurring in multiple overlapping contexts at multiple hierarchical levels, even adding ambiguity to a grammar is unlikely to yield a satisfactory representation of musical context. The use of context-sensitive grammars can address these problems to some extent but this also brings considerable additional difficulties in terms of efficiency and complexity. There are several attempts to model music using grammatical formalisms which add some degree of context sensitivity to context-free grammars without adding significantly to the complexity of the rewrite rules. An example is the Augmented Transition Network (ATN) which extends a recursive transition network (formally equivalent to a context-free grammar) by associating state transition arcs (rewrite rules) with procedures which perform the necessary contextual tests. *Cope* [25.91] describes the use of ATNs to rearrange harmonic, melodic, and rhythmic structures in EMI (experiments in musical intelligence). Another example is provided by the pattern grammars developed by *Kippen* and *Bel* [25.10] for modeling improvisation in North Indian tabla drumming.

Steedman [25.1, 72] has developed a categorial grammar (augmented context-free) to account for the harmonic structure of 12-bar blues, based on a theory of tonal harmony due to *Longuet-Higgins* [25.45, 46]. Although it is less ambitious than that of *Johnson-Laird* [25.85], this allows a more elegant description of improvisational competence since it does not rely on substitution into a previously prepared skeleton. However, in using the grammar to generate structural descriptions of blues chord progressions, Steedman had to introduce implicit meta-level conventions not explicit in the production rules of the grammar. The extent of context-sensitivity required to adequately model musical structure requires further investigation.

25.5 Discussion

This discussion of theoretical accounts of musical syntax raises several issues and questions which are driving current research:

1. How powerful a grammar do we need to represent the relationships present in musical structure? Are there examples of syntactic musical structures that require (mild) context sensitivity? How can multiple, polyphonic streams be represented by formal approaches?
2. How does musical syntax interact with other aspects of musical structure such as rhythm, metre, and timing? Or are these aspects also best explained using syntactic formalisms?
3. To what extent does real music, and listeners' perception of music, exhibit features of recursion, nonlocal dependencies, single or multiple center-embedding?
4. Which kinds of formal structures are listeners (musicians or nonmusicians) sensitive to?
5. Can such syntactic structures and relationships be learned, and if so, how and which kinds of predispositions need to be assumed as innate?

The power of a particular syntactic formalism is independent of whether it is probabilistic or deterministic. Probabilistic models have distinct advantages in terms of the subtlety with which they can capture structural dependencies for application in prediction, classification, parsing, and learnability/inference as well as in terms of robustness and graded grammaticality. For each of the model classes of the extended Chomsky hierarchy probabilistic counterparts have been proposed (e.g., finite-context grammars: *n*-gram models; regular grammars: Hidden Markov models; context-free grammars: probabilistic context-free grammars). These developments suggest as a general strategy that it may be beneficial to move from deterministic to probabilistic models for implementation and evaluation. It is important to note here that the Chomsky hierarchy is just one way of characterizing the power of grammatical formalisms but it does not necessarily lend itself naturally to every aspect of musical structure. Furthermore, as we noted above degrees of context-free character may be added to regular grammars and degrees of context-sensitivity added to context-free grammars.

While Markov and *n*-gram models are easily learned and are useful for prediction, they are barely capable of modeling more complex structures, nonlocal and hierarchical dependencies in music described above that are essential in musical implicative structure, stability, tension, and similarity. Conversely, more powerful types of syntactic formalisms are considerably more difficult to infer from data. It is not currently clear that we can develop one overarching theoretical stance that generalizes across musical styles and cultures. As in other areas of empirical musicology, the majority of research on musical syntax has focused on Western music and harmony in particular (with a few notable exceptions including [25.10, 12] and recent work by *Mavromatis* [25.92]). Different musical styles or traditions may emphasize different kinds of building block or show different degrees and kinds of complexity in their syntactic structure. Cross-cultural comparisons may have implications for evolutionary theories of music. While each process of inference requires predetermined (innate) assumptions about at least the search space and the structure of the model, it must be noted that cross-cultural universality by no means implies innateness of more than these assumptions.

Many of these questions are best addressed by implementing a computational theory as a computer model that embodies a particular theoretical stance on musical syntax and testing the model by comparing its behavior with human behavior. Modeling requires the theory to make all its assumptions explicit and permits the analysis of complex examples and large corpora. It is also possible to conduct a quantitative comparison of the behavior of a computational model with the behavior of listeners, allowing a rigorous empirical test of the theory as a psychologically plausible model of cognitive representation and processing of musical syntax. We address these points in detail in the following chapter, *Musical Syntax II*.

25.A Appendix: The Chomsky Hierarchy

Noam Chomsky introduced a containment hierarchy of four classes of formal grammar in terms of increasing restrictions placed on the form of valid rewrite rules [25.52]. A formal grammar consists of a set of nonterminal symbols (variables), terminal symbols (elements of the surface), production rules, and a starting symbol to derive productions. In the following description, $a \in T^*$ denotes a (possibly empty) sequence of terminal symbols, $A, B \in V$ denote nonterminal symbols, $\alpha \in (V \cup T)^+$ denotes a nonempty sequence of terminal and nonterminal symbols, and $\beta, \beta' \in (V \cup T)^*$ denote (possibly empty) sequences of terminal and nonterminal symbols. The difference between different formal grammars in the Chomsky hierarchy relates to different possible production rules.

Every grammar in the Chomsky hierarchy corresponds with an associated automaton: while formal grammars generate the string language, formal automata specify constraints on processing or generating mechanisms that characterize the formal language. Automata provide an equivalent characterization of formal languages as formal grammars.

25.A.1 Type 3 (Regular)

Grammars in this class feature restricted rules allowing only a single terminal symbol, optionally accompanied by a single nonterminal, on the right-hand side of their productions

$$A \to a$$
$$A \to aB \quad \text{(right linear grammar) or}$$
$$A \to Ba \quad \text{(left linear grammar)} .$$

Regular grammars generate all languages which can be recognized by a finite-state automaton, which requires no memory other than a representation of its current state.

It is essential to note that regular grammars are *not* equivalent to Markov models or k-factor languages (see Sect. 25.4.1 above).

25.A.2 Type 2 (Context Free)

Grammars in this class only restrict the left-hand side of their rewrite rules to a single nonterminal symbol – i. e., the right-hand side can be any string of nonterminal symbols

$$A \to \alpha .$$

The equivalent automata characterization of a context-free language is a nondeterministic pushdown automaton, which is an extension of finite-state automata that employs memory using a stack and state transitions may depend on the current state as well as the top symbol in the stack.

25.A.3 Type 1 (Context Sensitive)

Grammars in this class are restricted only in that there must be at least one nonterminal symbol on the left-hand side of the rewrite rule and the right-hand side must contain at least as many symbols as the left-hand side, i. e., string length increases monotonically in the production sequence.

$$\beta A \beta' \to \alpha, \quad |\beta A \beta'| \leq \alpha$$

Context-sensitive languages are equivalently characterized by a linear bounded automaton, that is a state-machine extended by a linear bounded random access memory band, whose transitions depend on the state, the symbol on the memory band.

25.A.4 Type 0 (Unrestricted)

Grammars in this class place no restrictions on their rewrite rules

$$\alpha \to \beta$$

and generate all languages which can be equivalently characterized by a universal Turing machine (the recursively enumerable languages), which is the same as the linear bounded automaton for context-sensitive languages without bounds on the memory tape.

References

25.1 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

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

25.3 M. Christiansen, N. Chater: Toward a connectionist model of recursion in human linguistic performance, Cogn. Sci. **23**, 157–205 (1999)

25.4 D. Sudnow: *Ways of the Hand: The Organization of Improvised Conduct* (MIT Press, Cambridge 1978)

25.5 M.T. Pearce, G.A. Wiggins: Auditory expectation: The information dynamics of music perception and cognition, Top. Cogn. Sci. **4**(4), 625–652 (2012), https://doi.org/10.1111/j.1756-8765.2012.01214.x
25.6 M. Rohrmeier, P. Rebuschat: Implicit learning and acquisition of music, Top. Cogn. Sci. **4**(4), 525–553 (2012), https://doi.org/10.1111/j.1756-8765.2012.01223.x
25.7 R.C. Berwick, A.D. Friederici, N. Chomsky, J.J. Bolhuis: Evolution, brain, and the nature of language, Trends Cogn. Sci. **17**(2), 91–100 (2013), https://doi.org/10.1016/j.tics.2012.12.002
25.8 E. Aldwell, C. Schachter: *Harmony & Voice Leading* (Thomson Schirmer, New York 2003)
25.9 I. Cross: Cognitive science and the cultural nature of music, Top. Cogn. Sci. **4**(4), 668–677 (2012)
25.10 J. Kippen, B. Bel: Modelling music with grammars. In: *Computer Representations and Models in Music*, ed. by A. Marsden, A. Pople (Academic Press, London 1992) pp. 207–238
25.11 S. Marcus: The Eastern Arab system of melodic modes: A case study of Maqam Bayyati. In: *The Garland Encyclopedia of World Music. The Middle East* (Routledge, New York 2003) pp. 33–44
25.12 D.R. Widdess: Aspects of form in North Indian ālāp and dhrupad. In: *Music and Tradition: Essays on Asian and Other Musics Presented to Laurence Picken* (Cambridge Univ. Press, Cambridge 1981) pp. 143–182
25.13 A. Sorace, F. Keller: Gradience in linguistic data, Lingua **115**(11), 1497–1524 (2005)
25.14 B. Aarts: *Fuzzy Grammar: A Reader* (Oxford Univ. Press, Oxford 2004)
25.15 B. Aarts: *Syntactic Gradience: The Nature of Grammatical Indeterminacy* (Oxford Univ. Press, Oxford 2007)
25.16 W. Caplin: *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven* (Oxford Univ. Press, New York, Oxford 1998)
25.17 E. Aldwell, C. Schachter: *Harmony and Voice Leading*, 2nd edn. (Harcourt Brace Jovanovich, San Diego 1989)
25.18 N. Chomsky: *Syntactic Structures* (Mouton, The Hague 1957)
25.19 N. Chomsky: *Aspects of the Theory of Syntax* (MIT Press, Cambridge 1965)
25.20 N. Chater: Reconciling simplicity and likelihood principles in perceptual organisation, Psychol. Res. **103**, 566–581 (1996)
25.21 N. Chater, P. Vitanyi: The generalized universal law of generalization, J. Math. Psychol. **47**, 346–369 (2003)
25.22 D.J.C. MacKay: *Information Theory, Inference and Learning Algorithms* (Cambridge Univ. Press, Cambridge 2003)
25.23 R. Cilibrasi, P.M.B. Vitanyi, R. de Wolf: Algorithmic clustering of music based on string compression, Comput. Music J. **28**, 49–67 (2004)
25.24 M.M. Marin, H. Leder: Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music, PLoS ONE **8**(8), e72412 (2013)
25.25 P.D. Grünwald: *The Minimum Description Length Principle* (MIT Press, Cambridge 2007)
25.26 A. Perfors, J.B. Tenenbaum, T. Regier: The learnability of abstract syntactic principles, Cognition **118**(3), 306–338 (2011)
25.27 C. Kemp, J.B. Tenenbaum: The discovery of structural form, Proc. Natl. Acad. Sci. **105**(31), 10687–10692 (2008)
25.28 P. Mavromatis: Minimum description length modelling of musical structure, J. Math. Music **3**(3), 117–136 (2009)
25.29 S. Kostka, D. Payne: *Tonal Harmony* (Alfred A. Knopf, New York 1984)
25.30 E. Narmour: *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model* (University of Chicago Press, Chicago, London 1992)
25.31 C.L. Krumhansl: *Cognitive Foundations of Musical Pitch* (Oxford Univ. Press, Oxford 1990)
25.32 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), Kobe*, ed. by K. Hirata, G. Tzanetakis, K. Yoshii (2009) pp. 549–554
25.33 S. Koelsch: Towards a neural basis of processing musical semantics, Phys. Life Rev. **8**(2), 89–105 (2011)
25.34 L.R. Slevc, A.D. Patel: Meaning in music and language: Three key differences: Comment on "Towards a neural basis of processing musical semantics" by Stefan Koelsch, Phys. Life Rev. **8**(2), 110–111 (2011)
25.35 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)
25.36 W.T. Fitch, B. Gingras: Multiple varieties of musical meaning: Comment on "Towards a neural basis of processing musical semantics" by Stefan Koelsch, Phys. Life Rev. **8**(2), 108–109 (2011)
25.37 F. Lerdahl: *Tonal Pitch Space* (Oxford Univ. Press, New York 2001)
25.38 M. Rohrmeier: Towards a generative syntax of tonal harmony, J. Math. Music **5**(1), 35–53 (2011)
25.39 M. Lehne, M. Rohrmeier, S. Koelsch: Tension-related activity in the orbitofrontal cortex and amygdala: An fMRI study with music, Soc. Cogn. Affect. Neurosci. **9**(10), 1515–1523 (2013)
25.40 M. Rohrmeier, W. Zuidema, G.A. Wiggins, C. Scharff: Principles of structure building in music, language and animal song, Phil. Trans. R. Soc. B (2015), https://doi.org/10.1098/rstb.2014.0097
25.41 G.J. Balzano: The pitch set as a level of description for studying musical pitch perception. In: *Music, Mind and Brain*, ed. by M. Clynes (Plenum, New York 1982) pp. 321–351
25.42 L. Euler: *Tentamen Novae Theoriae Musicae* (Academia Scientiae, St. Petersburg 1739), reprint: Broude Bros., New York 1968
25.43 G. Weber: *Versuch einer geordneten Theorie der Tonsetzkunst*, Vol. 1–4 (Schott, Mainz 1830)

25.44 J. Pressing: Cognitive isomorphisms between pitch and rhythm in world musics: West Africa, the Balkans and Western tonality, Stud. Music **17**, 38–61 (1983)
25.45 H.C. Longuet-Higgins: Letter to a musical friend, Music Rev. **23**, 244–248 (1962)
25.46 H.C. Longuet-Higgins: Second letter to a musical friend, Music Rev. **23**, 271–280 (1962)
25.47 R.N. Shepard: Structural representations of musical pitch. In: *Psychology of Music*, ed. by D. Deutsch (Academic Press, New York 1982) pp. 343–390
25.48 D. Tymoczko: *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice* (Oxford Univ. Press, Oxford 2011)
25.49 J. London: *Hearing in Time* (Oxford Univ. Press, Oxford 2004)
25.50 P. Janata, J.L. Birk, J.D. van Horn, M. Leman, B. Tillmann, J.J. Bharucha: The cortical topography of tonal structures underlying Western music, Science **298**(5601), 2167–2170 (2002)
25.51 G. Jäger, J. Rogers: Formal language theory: refining the Chomsky hierarchy, Philos. Trans. R. Soc. B **367**(1598), 1956–1970 (2012)
25.52 J.E. Hopcroft, J.D. Ullman: *Introduction to Automata Theory, Languages and Computation* (Addison-Wesley, Reading 1979)
25.53 G. Rozenberg, A. Salomaa (Eds.): *Handbook of Formal Languages* (Springer, New York 1997)
25.54 G. Wiggins: Computer models of (music) cognition. In: *Language and Music as Cognitive Systems*, ed. by P. Rebuschat, M. Rohrmeier, I. Cross, J. Hawkins (Oxford Univ. Press, Oxford 2012) pp. 169–188
25.55 M.A. Rohrmeier, S. Koelsch: Predictive information processing in music cogntion. A critical review, Int. J. Psychophysiol. **83**(2), 164–175 (2012)
25.56 T.C. Bell, J.G. Cleary, I.H. Witten: *Text Compression* (Prentice Hall, Englewood Cliffs 1990)
25.57 S. Bunton: *On-Line Stochastic Processes in Data Compression*, Doctoral Dissertation (University of Washington, Seattle 1996)
25.58 D. Huron: *Sweet Anticipation: Music and the Psychology of Expectation* (MIT Press, Cambridge 2006)
25.59 J.-P. Rameau: *Traite de l'harmonie reduite a ses principes naturels* (J.B.C. Ballard, Paris 1722)
25.60 W. Piston: *Harmony* (W.W.Norton, New York 1948)
25.61 R.O. Gjerdingen: Learning syntactically significant temporal patterns of chords, Neural Netw. **5**, 551–564 (1992)
25.62 V. Byros: Meyer's anvil: Revisiting the schema concept, Music Anal. **31**(3), 273–346 (2012)
25.63 V. Byros: Towards an "archaeology" of hearing: schemata and eighteenth-century consciousness, Musica Humana **1**(2), 235–306 (2009)
25.64 D. Conklin, I.H. Witten: Multiple viewpoint systems for music prediction, J. New Music Res. **24**(1), 51–73 (1995)
25.65 M.T. Pearce: *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition*, Doctoral Dissertation (Department of Computing, City University, London 2005)
25.66 L.R. Rabiner: A tutorial on Hidden Markov Models and selected applications in speech recognition, Proc. IEEE **77**(2), 257–285 (1989)
25.67 S. Kostka, D. Payne: *Tonal Harmony* (McGraw-Hill, New York 1995)
25.68 M. Rohrmeier: A generative grammar approach to diatonic harmonic structure. In: *Proc. 4th Sound Music Comput. Conf.*, ed. by Spyridis, Georgaki, Kouroupetroglou, Anagnostopoulou (2007) pp. 97–100
25.69 D.R. Hofstadter: *Goedel, Escher, Bach* (Basic Books, New York 1979)
25.70 I. Giblin: *Music and the Generative Enterprise: Situating a Generative Theory of Tonal Music in the Cognitive Sciences*, Doctoral Dissertation (University of New South Wales, Sydney 2008)
25.71 F. Lerdahl, R. Jackendoff: *A Generative Theory of Tonal Music* (MIT Press, Cambridge 1983)
25.72 M. Steedman: A generative grammar for jazz chord sequences, Music Percept **2**(1), 52–77 (1984)
25.73 H. Riemann: *Musikalische Syntaxis* (Breitkopf Härtel, Leipzig 1877)
25.74 T. Christensen: The Schichtenlehre of Hugo Riemann, Theory Only **6**(4), 37–44 (1982)
25.75 H. Schenker: *Der Freie Satz. Neue musikalische Theorien und Phantasien* (Margada, Liège 1935)
25.76 F. Lerdahl: Cognitive constraints on compositional systems. In: *Generative Processes in Music: The Psychology of Performance, Improvisation and Composition*, ed. by J.A. Sloboda (Clarendon, Oxford 1988) pp. 231–259
25.77 A. Keiler: Bernstein's "The Unanswered Question" and the problem of musical competence, Music. Q. **64**(2), 195–222 (1978)
25.78 A. Keiler: Two views of musical semiotics. In: *The Sign in Music and Literature*, ed. by W. Steiner (Univ. Texas Press, Austin 1981) pp. 138–168
25.79 D. Tidhar: *A Hierarchical and Deterministic Approach to Music Grammars and its Application to Unmeasured Preludes* (dissertation.de, Berlin 2005)
25.80 I. Deliège: Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules, Music Percept **4**(4), 325–360 (1987)
25.81 R. Jackendoff: *Foundations of Language – Brain, Meaning, Grammar, Evolution* (Oxford Univ. Press, Oxford 2003)
25.82 R. Jackendoff: A parallel architecture perspective on language processing, Brain Res **1146**, 2–22 (2007)
25.83 C. Roads: Grammars as representations for music. In: *Foundations of Computer Music*, ed. by C. Roads, J. Strawn (MIT Press, Cambridge 1985) pp. 403–442
25.84 J. Sundberg, B. Lindblom: Generative theories for describing musical structure. In: *Representing Musical Structure*, ed. by P. Howell, R. West, I. Cross (Academic Press, London 1991) pp. 245–272
25.85 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

25.86 S.M. Shieber: Evidence against the context-freeness of natural language. In: *The Formal Complexity of Natural Language*, ed. by W.J. Savitch, E. Bach, W. Marsh, G. Safran-Navah (Springer Netherlands, Dordrecht 1987) pp. 320–334
25.87 A.K. Joshi, K.V. Shanker, D. Weir: *The Convergence of Mildly Context-Sensitive Grammar Formalisms*. Technical Report No. MS- CIS-90-01 (Univ. of Pennsylvania, Department of Computer and Information Science 1990)
25.88 M. Steedman: *The Syntactic Process* (MIT Press, Cambridge 2001)
25.89 E.P. Stabler: Computational perspectives on minimalism. In: *Oxford Handbook of Linguistic Minimalism*, ed. by C. Boeckx (Oxford Univ. Press, Oxford 2011) pp. 617–643
25.90 J. Katz, D. Pesetsky: The Identity Thesis for Language and Music. http://ling.auf.net/lingbuzz/000959 (2010)
25.91 D. Cope: Computer modelling of musical intelligence in EMI, Comput. Music J. **16**(2), 69–83 (1992)
25.92 P. Mavromatis: A hidden Markov model of melody production in Greek church chant, Comput. Musicol. **14**, 93–112 (2005)