The Sweet Sounds of Syntax: Music, Language, and the Investigation of Hierarchical Processing

Lee Whitehorne*

University of Victoria

lwhitey@uvic.ca

Abstract

Language and music are uniquely human faculties, defined by a level of sophistication found only in our species. The ability to productively combine contrastive units of sound, namely words in language and notes in music, underlies much of the vast communicative and expressive capacities of these systems. Though the intrinsic rules of syntax in language and music differ in many regards, they both lead to the construction of complex hierarchies of interconnected, functional units. Much research has examined the overlap, distinction, and general neuropsychological nature of syntax in language and music but, in comparison to the psycholinguistic study of sentence processing, musical structure has been regarded at a coarse level of detail, especially in terms of hierarchical dependencies. The current research synthesizes recent ideas from the fields of generative music theory, linguistic syntax, and neurolinguistics to outline a more detailed, hierarchy-based methodology for investigating the brain's processing of structures in music.

Keywords: music cognition; music perception; syntax; generative grammar; structural processing

Language and music are highly sophisticated and uniquely human faculties (Jackendoff, 2009; Lerdahl & Jackendoff, 1983; Patel, 2007). One of the fundamental elements that distinguishes human language from other forms of animal communication is the near-endless human capacity to recombine units of meaning and grammatical function in novel ways to meet our communicative and expressive needs. We group words into abstract, conceptual categories—our nouns, verbs, adjectives, and other parts of speech—and assemble them based on implicit rules known by the native speakers of a given language. Similar structural patterns have also been noted in music, though the communicative ability, sonic properties, and categorization within that domain are vastly different from those of language. Put in general terms, both domains feature the combination of discrete sound units into much larger forms.

Of course, relationships between units in either domain are not of a purely sequential nature. Connections are instead formed between events of relative structural importance or function, creating complex hierarchies based on rules known implicitly by the listener (Jackendoff, 2009; Lerdahl & Jackendoff, 1983; Patel, 2007). As demonstrated in Figure 1, these hierarchical dependencies can occur between non-adjacent units in both language and music, even at large distances, and are essential for understanding incompatibilities in a given sequence. A ready analogue to sentence structure in music comes from tonal relations—systems of musical keys, modes, and harmonies, as described by Western musical theory—found across expansive musical passages or even whole pieces. The examples given in Figure 1 illustrate how the final unit of each sequence (the main verb in the linguistic examples and the final chord in the musical ones) is dependent on a unit at the beginning of the sequence, rather than those directly adjacent to it. In the linguistic examples, this unit is

the sentence subject "The dogs." In the musical examples, the first chord acts as the tonal centre, the most stable or consonant harmony (or single pitch, in other cases) in the sequence which all other events are related back to. Figure 1 also demonstrates how such relationships in either domain can be represented through the use of tree diagrams, with more structurally important connections effectively occurring higher in a given tree, though the intricacies of how these given analyses were reached will not be made explicit at this point. How, then, does the brain parse acoustic input and assemble that information into the proposed hierarchies? The purpose of the research described here is to look deeper into the structures underlying music and language, and to develop new ways of investigating their cognitive foundations.

Figure 1: *Long-distance dependencies in language and music. A noun in the subject separated from a compatible (above) or incompatible (below) verb by a relative clause; analogous musical sequences with in-key (above) and out-of-key (below) final chords.*

The generative power of syntax, responsible for much of the generative power of language itself, has already invited much empirical inquiry that has examined the exact nature of what mechanisms and processes in the human brain allow such a feature to manifest. By varying the content of both target sentences and musical passages, for example, experimental studies have observed the effects of structural differences on neural processing (Maidhof & Koelsch, 2011). Results have suggested facilitation of behavioural responses from structural priming in congruent sequences (i.e., when words or chords fit into their preceding contexts) and response inhibition from structural violations (i.e., semantically unrelated words or dissonant chords) (Wright & Garrett, 1984). Some have also documented the neural activity related to those events (Koelsch, Gunter, Friederici, & Schröger, 2000; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Loui, Grent-'t-Jong, Torpey, & Woldorff, 2005).

Though these findings inform much of what we understand about syntax and the brain, their methodologies typically adopt linear representations of the structures in question, falling short of the explicit degree of hierarchical detail defined by theories of both linguistic syntax and musical structure, not to mention neuropsychological experiments that have tested said models within the language domain (Ding, Melloni, Zhang, Tian, & Poeppel, 2015; Lerdahl & Jackendoff, 1983). This article outlines a movement toward hierarchy-centred methodologies for studying structural processing of music. Previous approaches to musical structure will be reviewed from both the experimental and theoretical literature, illustrating some useful parallels from the investigation of linguistic syntax. From those foundations, a new methodology for defining musical hierarchies is

outlined and demonstrated, resulting in the creation of a novel, hierarchy-based stimulus paradigm. Potential experimental applications for this paradigm are then outlined, as well as further possible directions for this vein of research.

Past Investigations of Hierarchical Processing

Musical Stucture

Previous attempts at studying the processing of hierarchies in music have focused on relations to a tonal centre, or whether a musical event is heard as belonging to an overarching key or not (Koelsch et al., 2000; Koelsch, Fritz, Schulz, Alsop, & Schlaug, 2005; Koelsch et al., 2013; Loui et al., 2005). Many have used stimulus paradigms composed of five-chord sequences, manipulating target chords in terms of their harmonic congruence (i.e., chords heard as more consonant or dissonant within a given context) (Koelsch et al., 2000; Koelsch et al., 2005; Loui et al., 2005). Long-distance dependencies have also been examined on a larger scale, through modulating sections of a complete musical piece away from an established tonal area (Koelsch et al., 2013). Both stimulus approaches have revealed predictable neural responses to unexpected (or less stable) harmonic events, providing some physiological indicators of hierarchical dependencies. For example, a number of studies have observed brain activity, an early right anterior negativity (ERAN), in response to structural violations in music (i.e., incongruent chords) that resembles an analogous early left anterior negativity (ELAN) found in language processing, though with slightly different timing and localization (Koelsch et al., 2005; Koelsch et al., 2013; Loui et al., 2005; Maidhof & Koelsch, 2011). Acknowledging that these studies provide some evidence for non-adjacent dependencies between events in music (i.e., belonging to a shared musical key), they fall short of defining these structural relationships in much detail, especially when compared to linguistic approaches described in the next section (Madell & Hébert, 2008).

Looking to Linguistic Syntax

A notable approach to the issue of hierarchy-building in language looked at online processing of these relationships over time, comparing patterns of brain activity across the duration of differently structured linguistic expressions (Ding et al., 2015). Mandarin words were grouped into intermediate levels of syntactic structure (noun and verb phrases—NPs and VPs) and combined to form four-word sentences. These sequences were then presented auditorily to participants without any audible pauses or breaks between the groupings. The peaks of brain activity observed in participants were ultimately correlated with the time course of those phrase-level syntactic constituents, not just at the word and sentence levels. These results suggest a neural basis for the fine-grained hierarchies that have long been central to theories of linguistic syntax. This approach was also particularly innovative for its investigation of hierarchical processing using grammatical linguistic sequences, in contrast to violation-based approaches analogous to those outlined in the previous section on musical processing (Koelsch et al., 2000; Koelsch et al., 2005; Loui et al., 2005; Maidhof & Koelsch, 2011). In other experiments, lexical decision tasks have also demonstrated the effect of different types of phrase-level contexts on the judgment time of a target word (Wright & Garrett, 1984). By manipulating grammatical intermediate-level constructions and not simply individual events, these approaches suggest some potential directions for exploring analogous types of structure in music.

The adaptation of these psycholinguistic approaches to a musical domain does raise some key issues, however. For one, the grammatical categories that define phrase-level constituents in language have no clear parallel in music (Jackendoff, 2009; Lerdahl & Jackendoff, 1983; Patel, 2007). Though

recent research suggests that listeners may group similar chords into functional categories based on the predictable contexts they appear in, it remains unclear whether nouns and verbs have counterparts in music (Goldman, Jackson, & Sajda, 2018). Additionally, though a word placed unexpectedly in the context of a sentence may categorically violate our intrinsic grammatical expectations, even highly-marked deviants within a musical sequence (e.g., unexpected or unfamiliar harmonies) may find resolution through integration with the following context (Lerdahl & Jackendoff, 1983). This illustrates the importance of relative stability between events in music, in contrast to the more categorical rules of grammaticality found in language. Acknowledging these differences, neural imaging and electrophysiological studies have still demonstrated that our brain at least responds to differences of structural congruency in analogous—however distinct—ways in both music and language (Koelsch et al., 2005; Koelsch et al., 2013; Loui et al., 2005; Maidhof & Koelsch, 2011; Patel, 2007; Rogalsky, Rong, Saberi, & Hickok, 2011). Therefore, in order to better examine the hierarchical nature of musical structure, those structural relationships must be precisely defined, yet understood in cognitively realistic terms appropriate to the musical domain.

A Generative Approach to Musical Hierarchies

Looking toward the theoretical literature, Lerdahl and Jackendoff's *Generative Theory of Tonal Music* (1983), or GTTM, describes a unique and rigorous approach to analyzing musical structures, applying concepts from the fields of linguistic syntax, musical theory, and cognitive science to formulate a universally applicable theory of musical grammar. In addition to serving as a tool for analysis, however, this theory could help guide new experimental methodologies for investigating hierarchical processing.

Building in part on ideas developed by the early twentieth-century music theorist Heinrich Schenker, GTTM culminates in a system for reducing musical works to their key prolongational relationships, the implied continued "hearing" of important and stable musical events while the musical surface (the actual notes being heard) morphs and diverges. Prolongation can be observed as the sense of "tensing" and "relaxing" a listener experiences when they listen to a piece of music or, alternately, the expectation of certain important musical events to return, and the eventual resolution (or lack thereof) when that return occurs (or fails to do so). This sense of expectation and resolution (or diversion) between non-adjacent constituents can also be found in language processing and, most importantly, found rooted in syntactic structure (Patel, 2007; Wright & Garrett, 1984). Transitivity in verbs, for example, can strongly suggest the continuation of a sentence, as found in a sequence like "He gave his sister...". Use of the verb "gave" in this case demands both a direct and indirect object will be present, producing an effect of anticipation in a listener (or reader) as the sentence unfolds a second object or trails off, incomplete. Prolongation in music is ultimately described by Lerdahl and Jackendoff (1983) through explicit hierarchical structures, with all musical events related through recursive branching from events of greater stability, known as prolongational "heads"; less stable events are considered to be "elaborations" of the prolongational heads they branch from. This principle of "headedness" within the prolongational analytical system creates another key parallel with theories of linguistic syntax, many of which also assume head-based hierarchies (Lerdahl & Jackendoff, 1983; Patel, 2007). Prolongational heads will be discussed in greater detail in later sections.

GTTM defines a set of generative rules for analyzing prolongational structures, based on principles of well-formedness (requirements underpinning their analytical approach) and principles of preference (inherent tendencies of the listener to prefer certain potential analyses over others). The well-formedness rules specify all of the possible analyses that could be applied to specific musical passages, without consideration of which analysis would be deemed most correct based on a listener's

implicit musical knowledge. As a brief clarification, Lerdahl and Jackendoff's (1983) use of the term "well-formedness" in this manner may unfortunately cause some confusion, as the "well-formedness" of a linguistic element generally refers to how well it obeys the specific grammatical rules of a given language. Idiomatic musical features are instead represented through the preference rules, further emphasizing the more-gradient and less-categorical nature of musical grammaticality, as discussed earlier. Preference rules independently take into account how musical events are perceptually grouped together, as well as how regular metrical patterns (i.e., beats, pulses) are analyzed in the music, integrating these systems together to optimally describe the relative structural importance of the musical piece, referred to in GTTM as time-span segmentation and reduction. Perceptual groupings, metrical patterns, and time-span structures are all represented as explicit hierarchical relationships, defined through Lerdahl and Jackendoff's generative rules, though only the time-span and prolongational systems form headed structures in a way that parallels linguistic syntax. As the ultimate result of prolongational reduction, all musical events in a piece are proposed to be related through either strong prolongation (an exact repetition of an event's harmonic content), weak prolongation (repetition of pitch content but with harmonic roots—the bass and melody notes—on different pitches within the harmony, known as inversion in Western music theory) or progression (changes in harmonic content and different pitch classes). These different types of elaboration are illustrated in Figure 2.
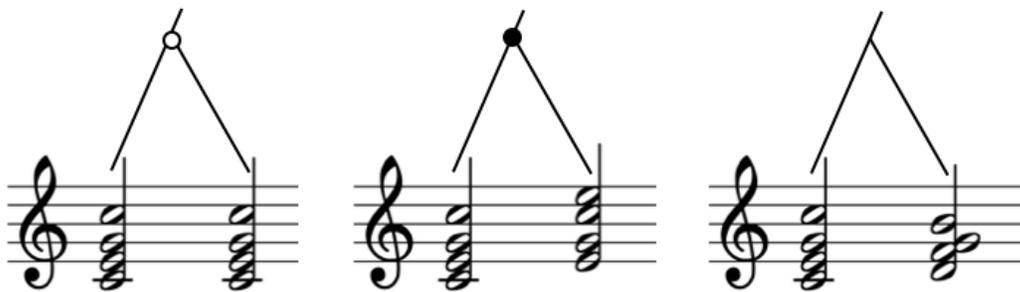

Figure 2: *Three possible elaborations of a C major chord: a strong prolongation (left, indicated with an open circle node), a weak prolongation (centre, indicated with a closed circle node), and a progression (right).*

Much as with linguistic syntax, the resulting analyses can be represented using both linear and tree diagrams. This fine-grained approach to analyzing hierarchical structures provides the necessary theoretical background for developing a new experimental approach, which is outlined in the following section.

Developing a New Paradigm

GTTM as a Framework for Constructing Hierarchical Structures

To investigate the nature of hierarchical processing of music, the structures of any experimental stimuli should be explicitly defined and then manipulated. Lerdahl and Jackendoff's (1983) systems of time-span and prolongational reduction provide a clear, rule-based framework for creating these structures. In fact, both systems mutually influence the analyses of the other, as will become apparent later in this section. For the purposes of this article, the aspects of these systems relevant to analyzing short, isolated musical sequences (such as those used in experimental settings) will now be cursorily defined. The interpretation of the rules for this application also assumes features

idiomatic to the Western classical music tradition, though Lerdahl and Jackendoff make clear where rules may vary between different musical traditions of the world. The choice of focusing on the Western classical idiom here is based on overwhelming precedent in the experimental literature and familiarity for the author. All musical examples and figures that follow (including the circle of fifths diagram) were composed or constructed by the author for illustrative purposes.

The Rules of GTTM

Before establishing prolongational relationships between events, the time-span importance of the musical events in question must be considered first. This analytical system takes input mostly from the independent analyses of grouping and metrical structures, the details of which will not be elaborated on here but are described thoroughly within GTTM. Time-span reduction involves "the segmentation of a piece into rhythmic units within which relative structural importance of pitch-events can be determined" (Lerdahl & Jackendoff, 1983). Within any given time-span, one event (or one smaller time-span contained within the time-span in question) is chosen as the time-span "head," the most structurally important event. A time-span head is chosen according to the following preference rules (time-span reduction preference rules, or TSRPRs), paraphrased from Lerdahl and Jackendoff's own proposals:

1. Prefer a head on a strong beat.

2. Prefer a head that is more intrinsically stable and/or closely related to the local tonic (most stable harmonic event).

3. Weakly prefer a head with a higher melody or lower bass note.

4. If two time-spans appear to be parallel (comprised of very similar melodic, rhythmic, and/or structural patterns), prefer to assign them parallel heads.

5. Prefer a head that results in a more stable metrical structure.

6. Prefer a head that results in a more stable prolongational structure.

7. If a sequence of events forms a cadence at the end of the time-span, prefer the cadence to be labelled as the head.

8. If the time-span in question is at the beginning of a larger time-span, prefer a head that is close to the beginning of the time-span.

One additional preference rule, TSRPR 9, is defined in GTTM, though it is only relevant at the level of a complete musical piece and therefore not discussed here.

When constructing experimental stimuli, these factors operate in a number of ways (relevant rules are indicated in parentheses). For short sequences of chords, little context will be available to establish metrical regularity. As illustrated in Figure 3, listeners tend to group beats (i.e., chords) in twos or threes, dependent on the relative harmonic stability of the events (TSRPR 2) and the time-span analysis of the preceding context (TSRPRs 4 and 5); the first beat of each group will also serve as a beat at the higher level of metrical structure (a "strong" beat) (Lerdahl & Jackendoff, 1983).

The first beat of each chord sequence will be preferred as the time-span head (TSRPRs 1, 5, and 8), though this can be subverted by decreasing its harmonic stability (TSRPR 2), as shown in Figure 4. In general, root position triads are the most intrinsically consonant, becoming less consonant in different inversions and/or with the addition of extra pitches (e.g., adding the seventh
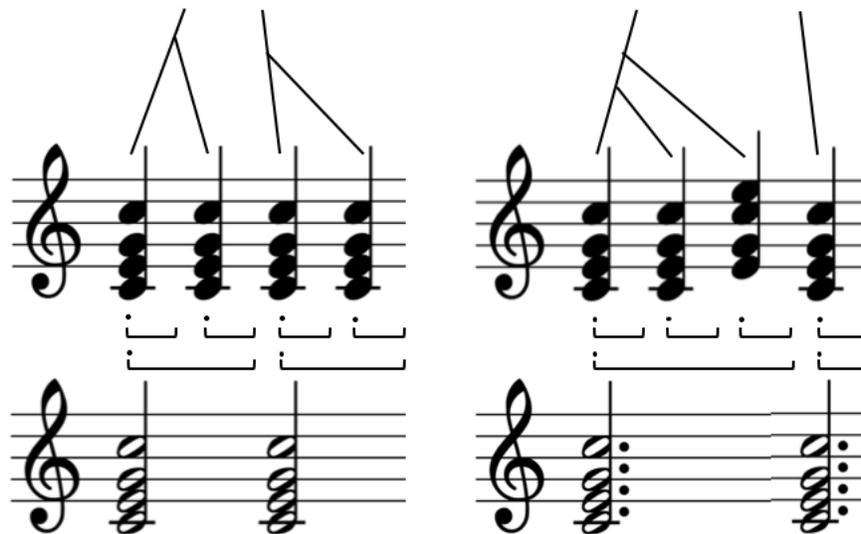
Figure 3: *Time-span reduction of two simple chord sequences (top staves), represented using tree diagram notation (above) and musical staff notation (bottom staves). Dots and brackets beneath the top staff represent different levels of metrical analysis (beats) and time-span segmentation, respectively.*
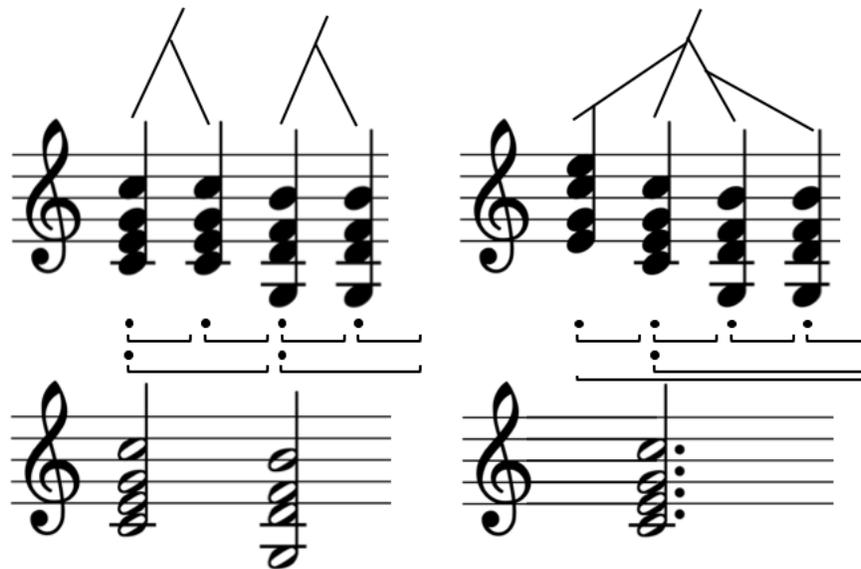
Figure 4: *Time-span reduction of two simple chord sequences (top staves). Though the first chord of each sequence contains the same pitch classes, the less stable inversion of that chord in the rightmost sequence leads to a markedly different time-span reduction.*

to a dominant V chord). With that in consideration, the metrical analysis (and, consequently, the time-span analysis) can therefore be strongly influenced by the placement of these maximally consonant chords, as demonstrated by the contrast between the first two sequences in Figure 5. Conversely, a strongly consonant chord placed directly after an identical chord at the beginning of a sequence will likely have less time-span importance due to the preceding chord sounding like a stronger beat (TSRPR 1), as shown by the analysis of the third sequence in Figure 5. Finally, any cadential sequence at the end of a chord sequence, especially a dominant (V) to tonic (I) progression, will collectively be a more important time-span (TSRPR 7), as illustrated by Figure 6. TSRPR 6, in this application, is automatically satisfied by the intentional selection of maximally stable prolongational structures to support an experimental design.
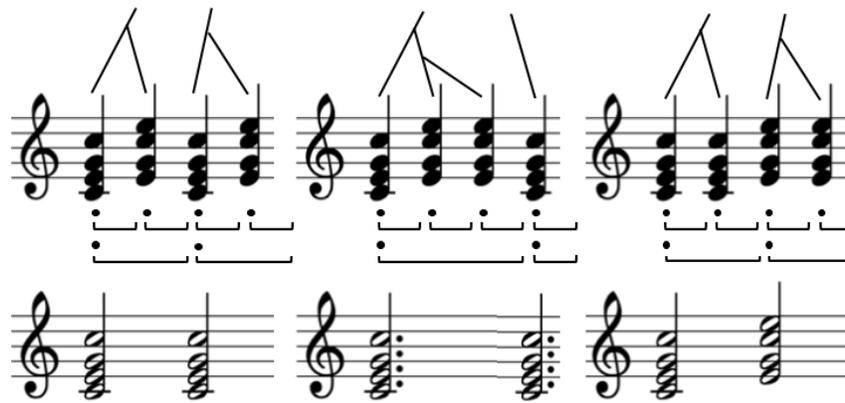

Figure 5: *Time-span reductions resulting from differing placements of root position triads.*
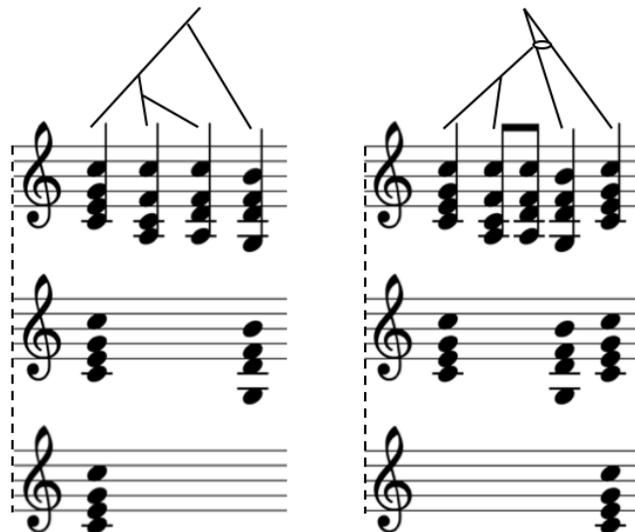

Figure 6: *Two levels of time-span reduction for two similar chord sequences, the rightmost ending in a cadence. Note the difference in analysis due to retention of the cadence in the time-span reduction.*

Like time-span reduction, prolongational reduction also segments a musical sequence, but into hierarchically-related regions that represent either an overall "tensing" or "relaxing," "strongly influenced by the relative importance of events in the time-span reduction" (Lerdahl & Jackendoff, 1983). A prolongational head is chosen for each region, this time representing the most prolongationally stable event in that region. In a tree diagram representation of a prolongational reduction, an increase in tension over time is shown by right-branching elaborations, and a decrease in tension over time is shown by left-branching elaborations. The proposed prolongational reduction preference rules (PRPRs) for choosing a prolongational head are paraphrased as follows:

1. Prefer a head which has a relatively high time-span importance.

2. Prefer elaborations of more stable events within the same time-span, rather than across different time-spans.

3. Prefer elaborations that form maximally stable connections with more stable events.

    3.1. Branching condition (see Figure 7):

        a. Right-branching elaborations are most stable if strong prolongations (exact repetitions) and least stable if progressions (different chords).

        b. Left-branching elaborations are most stable if progressions, least stable if strong prolongations.

    3.2. Connections between events are more stable if common pitch collections are involved or implied (see Figure 8).

    3.3. Melodic condition (see Figure 9):

        a. Connections are more stable if the melodic interval between them is smaller.

        b. Ascending melodies are more stable as right-branching elaborations; descending melodies are more stable as left-branching elaborations.

    3.4. Harmonic condition (according to Western classical common practice) (see Figure 10):

        a. a. Connections are more stable if chord roots are closer together on the circle of fifths (i.e., the number of stacked perfect fifth intervals needed to reach one pitch class from another, shown in Figure 11).

        b. Progressions ascending the circle of fifths are more stable as right-branching elaborations; progressions descending the circle of fifths are more stable as left-branching elaborations.

4. Prefer elaborations of more prolongationally stable heads (see Figure 12).

5. Prefer parallel prolongational analyses for parallel sequences (those comprised of very similar melodic, rhythmic, and/or structural patterns).
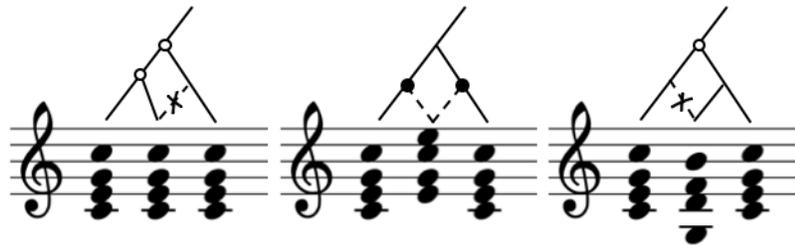
Figure 7: *Illustration of PRPR 3.1 (Branching Condition), which prefers the given prolongation analyses based on the observed types of elaboration (i.e., prolongation vs. progression). The centre analysis remains ambiguous without more context.*
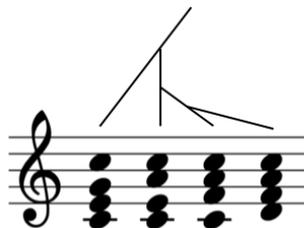

Figure 8: *Illustration of PRPR 3.2, which prefers connections between events that share a common pitch collection (pitches C and E between chords 1 and 2, pitches A and C between chords 2 and 3, and pitches F, A, and C between chords 3 and 4).*
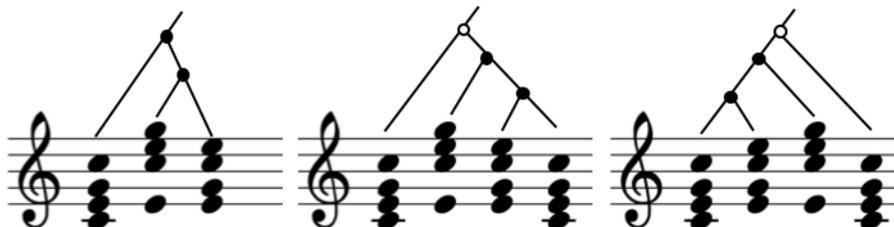

Figure 9: *Illustration of PRPR 3.3 (Melodic Condition), which prefers the given prolongational analyses based on melodic direction and distance instead of harmonic factors.*
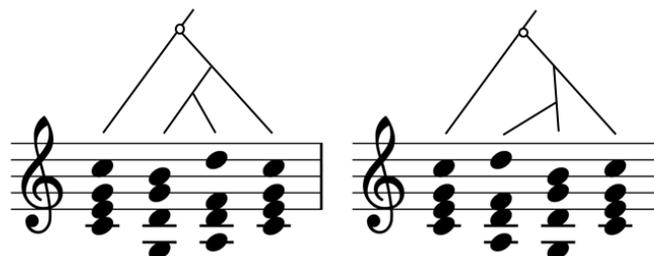

Figure 10: *Illustration of PRPR 3.4 (Harmonic Condition) which prefers the given prolongational analyses based on direction and distance between the middle two chord roots along the circle of fifths.*
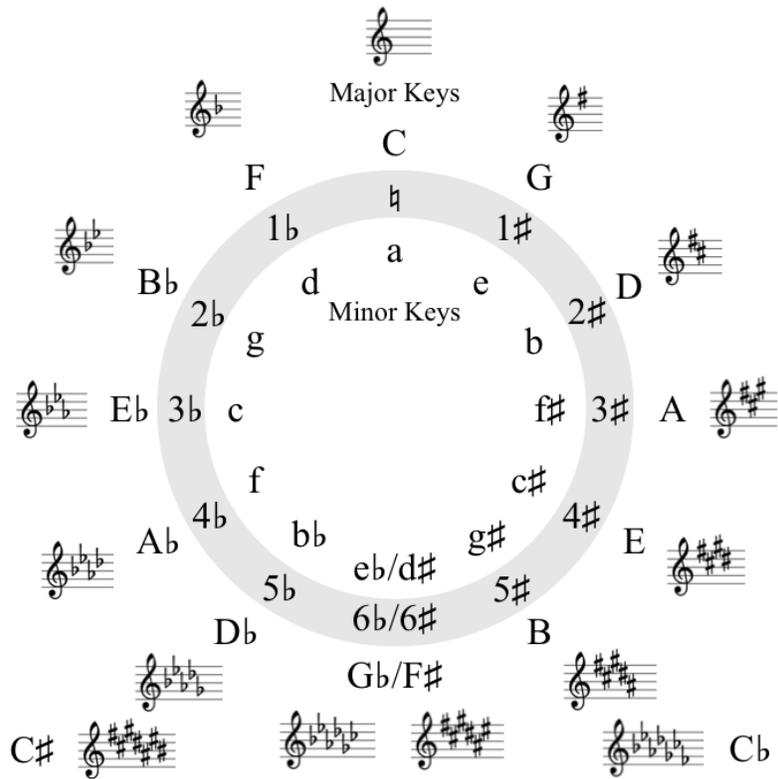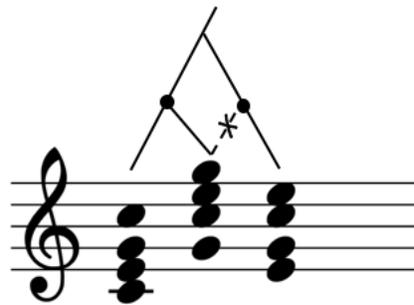
Figure 11: *The circle of fifths.*

Figure 12: *Prolongational tree for a short chord sequence. Though the last two chords are related through weak prolongation, they are both direct elaborations of the first chord due to its stability being greatest.*

A sixth rule and so-called Interaction Principle are also described by Lerdahl and Jackendoff, but they are not relevant to the discussion here, based on the scope of the musical structures in question.

For experimental design purposes, consideration of time-span importance is therefore quite important for creating a viable prolongational analysis (PRPR 1 and 2). The other factors presented above can be overridden if presented in a certain time-span context, as illustrated by Figure 13. Otherwise, PRPRs 1 through 5 are somewhat independent and self explanatory. Most structural manipulations are therefore dependent on relating events through weighting the various branching preferences. A specific application of these rules is described in the following section.
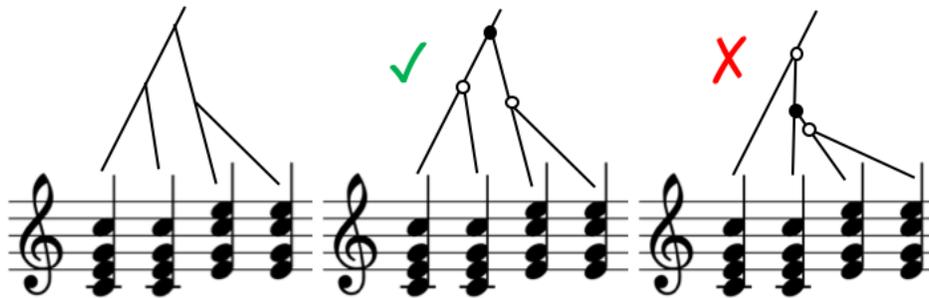

Figure 13: *Time-span tree (left) and two prolongational trees (centre, right) for a chord sequence. The Branching Condition of PRPR 3 alone would suggest the second prolongational analysis, due to the relative stability of strong versus weak prolongations; consideration of time-span importance ultimately leads to adopting the first analysis, however.*

An Example Paradigm

The current project involved developing a new stimulus paradigm for the experimental study of structural processing in music. Sequences of four chords were composed following the four-word stimuli used by Ding et al. (2015) for their neural investigation of sentence processing, as well as numerous five-chord paradigms used to investigate musical structure (Koelsch et al., 2000; Koelsch et al., 2005; Loui et al., 2005). The first chord of each sequence functioned as the main prolongational head, asserted by placing a root position major chord in that position of unambiguously high time-span importance.

Each sequence varied in its underlying prolongational structure, representing every hierarchy combinatorially possible for that number of musical events with the first chord as prolongational head. As prescribed by GTTM, this results in a total of twelve structures, without considering the different types of possible elaboration (prolongation vs. progression). With the first chord of each block serving as the prolongational head, the most stable event in the hierarchy, every other chord therefore acts as a recursive elaboration of that event. Note that each chord sequence is to be presented audibly in an experimental setting with uniform duration, intensity, timbre, and articulation, minimizing the confounding impact of those elements on the grouping and metrical analyses of a given passage, which affects its time-span and, consequently, prolongational reductions. The prolongational relationships within this paradigm are therefore based primarily on pitch collection (whether notes are shared between two chords), register (inversion of harmonic roots and octave displacement), harmonic distance (based on the circle of fifths), and melodic conditions. For the current scope of this project, different stimuli were created for each type of elaboration possible for the final (target) chord, while the other chords were only elaborated to minimize prolongational ambiguity and held constant

when possible. Nine additional sequences were added to represent certain hierarchies, varying their musical surface material in order to facilitate experimental counterbalancing. This resulted in a total of 45 blocks for the current paradigm, shown in Figure 14 with prolongational analyses shown for each structure.

Moving Forward

Experimental Approaches

As a next stage of the work reported here, an experiment has been designed to test the efficacy of this new hierarchy-based stimulus paradigm within a behavioural setting. Participants will be presented with a block of the current four-chord sequences, representing a set of contrasting prolongational structures, and asked to judge whether two target chords of each sequence are the same or different as quickly as possible. This judgment task directly addresses the prolongational relationship between those two chords, but is also anticipated to expose priming effects from hierarchical dependencies present within the preceding context as well. The expectation created by these constructed prolongational relationships is hypothesized to affect judgment task reaction times, analogous to what has been observed in psycholinguistic lexical decision studies (Wright & Garrett, 1984). Other four-chord paradigms could easily be developed for these applications as well, creating new hierarchies using the methodology described in the previous section. Long-distance dependencies could also be investigated by expanding these principles to longer musical sequences, potentially using eye tracking of sight-reading performers as a novel experimental task (Madell & Hébert, 2008).

A further application of these stimuli may be found in neural tracking experiments, investigating the processing of hierarchy-building in music. Though much prior research has identified and studied the brain's event-related potentials (ERPs) associated with unexpected harmonic events in a musical sequence (Koelsch et al., 2000; Koelsch et al., 2013; Loui et al., 2005; Maidhof & Koelsch, 2011), the time-course approach taken by Ding et al. (2015) serves as a promising framework for investigating different levels of musical hierarchy in the brain. The explicit structural dependencies in this new paradigm allow for a more precise manipulation of the phenomena to be tested and can help tease apart the different structural factors that together form our perception.

Non-Western/Classical Musical Idioms

The paradigm designed for this project falls into a common but unfortunate trend found throughout music cognition and perception research: an exclusive focus on a musical idiom of the Western European common-practice (classical) tradition (Jackendoff, 2009; Lerdahl & Jackendoff, 1983; Patel, 2007). Though the neural mechanisms for processing musical structures may be shared across the human species, the various elements that comprise music—pitch, rhythm, timbre, and more—play different structural roles across cultures and traditions. The massive importance of harmony in Western music, for example, is actually quite unique among the world's musics. Using the non-idiom-specific rules and abstract structural patterns described in GTTM combined with the methodology developed here, however, it may be possible to develop new paradigms based on the musical vocabularies of other traditions. From there, we can better investigate how hierarchical structure is processed universally, as well as what neuropsychological effects different levels of familiarity with a musical idiom might create.
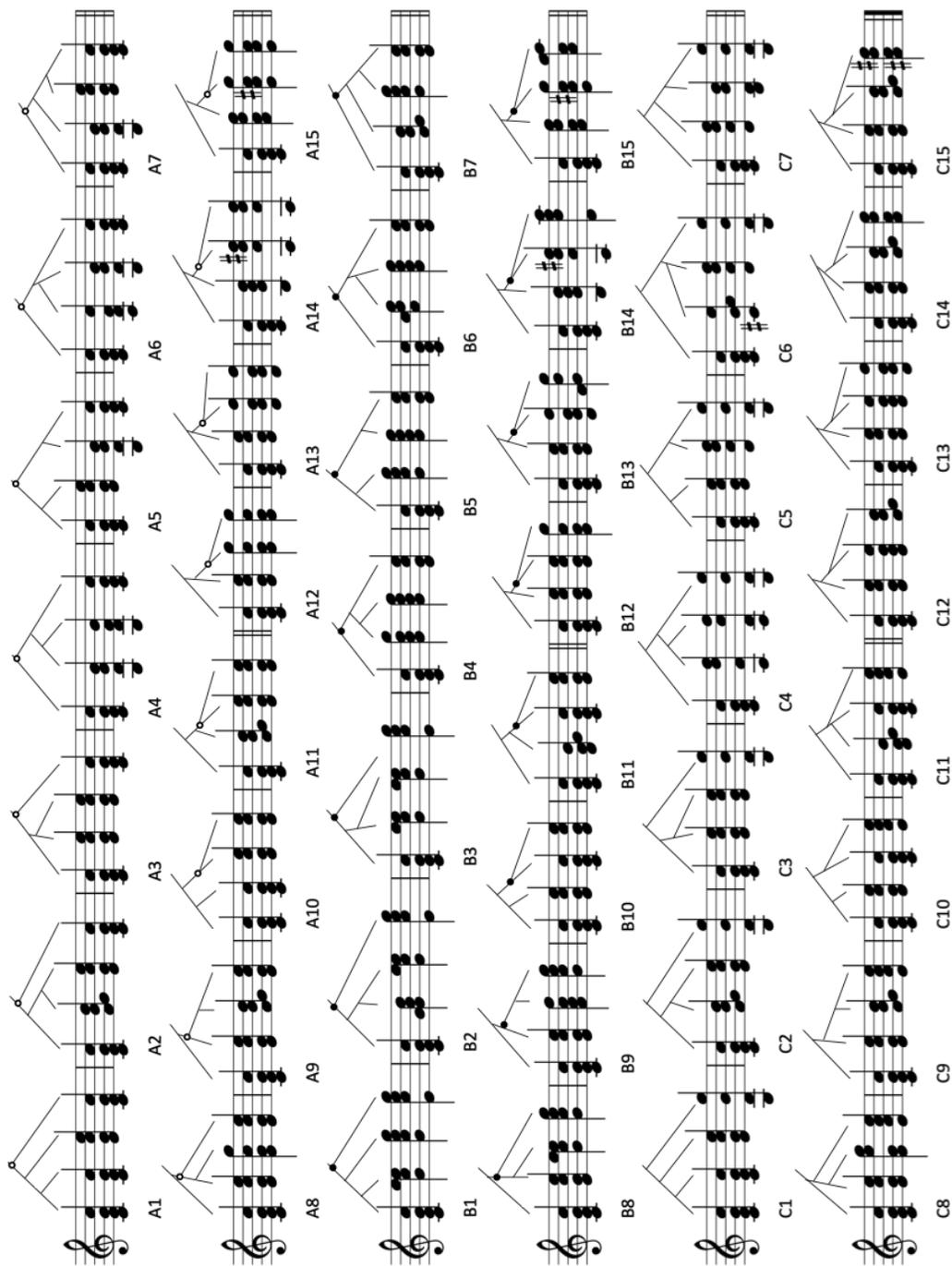
Figure 14: *An example experimental stimulus paradigm, with prolongational reduction tree diagrams. Open circles at branching nodes indicate strong prolongation of final (target) chord (top rows); closed circles indicate weak prolongations (middle rows) and bare nodes indicate progressions (bottom rows). Other elaboration types are not notated.*

Conclusion

The parallel yet divergent natures of music and language create a rich foil for scientific comparison. Beyond examining the cognitive and neurological underpinnings of these human faculties, however, researchers can also learn much from the theoretical and methodological approaches used in each opposing domain. By using a linguistically-informed cognitive theory of music and adapting a neurolinguistic experimental methodology for the musical domain, this article proposes new directions toward investigating structural processing in music, with a focus on how the brain constructs complex hierarchies from a stream of musical input. Building on the basic framework outlined here, further approaches could better explore how hierarchical structures are processed in both music and language, how expertise in different musical traditions influences these systems, and how exactly the brain integrates the multitude of elements that form these complex constructions.

References

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience, 19*, 158–164. https://doi.org/10.1038/no.4186

Goldman, A., Jackson, T., & Sajda, P. (2018). Improvisation experience predicts how musicians categorize musical structures. *Psychology of Music, 0*(0), 1–17. https://doi.org/10.1177/0305735618779444

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception: An Interdisciplinary Journal, 26*(3), 195–204. http://doi.org/10.1525/mp.2009.26.3.195

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience, 12*(3), 520–541. http://doi.org/10.1162/089892900562183

Koelsch, S., Fritz, T., Schulz, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage, 25*(4), 1068–1076. http://doi.org/10.1016/j.neuroimage.2004.12.050

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America, 110*(38), 15443–15448. https://doi.org/10.1073/pnas.1300272110

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

Loui, P., Grent-'t-Jong, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Cognitive Brain Research, 25*(3), 678–687. http://doi.org/10.1016/j.cogbrainres.2005.08.019

Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience, 23*(9), 2252–2267. https://doi.org/10.1162/jocn.2010.21542

Madell, J., & Hébert, S. (2008). Eye movements and music reading: Where do we look next? *Music Perception, 26*(2), 157–170. https://doi.org/10.1525/mp.2008.26.2.157

Patel, A. D. (2007). *Music, language and the brain.* New York, NY: Oxford University Press.

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *Journal of Neuroscience, 31*(10), 3843–3852. https:doi.org/10.1523/JNEUROSCI.4515-10.2011

Wright, B., & Garrett, M. (1984). Lexical decision in sentences: Effects of syntactic structure. *Memory & Cognition, 12*(1), 31–45. https://doi.org/10.3758/BF03196995