Throughout my journey learning to play the cello, I have sought to describe my repertoire in the most personal ways possible, in order to connect more deeply with the music and therefore perform the pieces more convincingly. However, as I explored more recently written works outside of the “standard” repertoire, I encountered phrase structures that I did not understand how to shape effectively and established forms of analysis offered little aid. My teachers’ suggestions included exacerbating the dynamics, thinking of lines as gestures, and focusing on tone color. Although these ideas helped tremendously, I sought a more systematic and quantifiable way to describe phrasing. As an educator, what would I tell my students to help them understand and perform such music?

Performing Baroque, classical, or romantic music without an understanding of the meter seems absurd, yet little exists to explain the temporal structures that operate outside of conventional meter. I was invigorated to develop an analytical framework upon hearing the music of Lee Hyla (1952–2014), a composer born in the United States of America and active in New York City, Boston, and Chicago. I was immediately struck by the energy and intensity in his works, particularly in sections in which I could not hear a steady pulse, and had strong emotional reactions to his music despite not immediately understanding how it worked. I was most taken by Hyla’s *Dream of Innocent III* (DOI3) (1987) for amplified cello, piano, and percussion, which underscores formal divisions with differing temporal aesthetics. The score further piqued my interest with markings such as *fast and brutal* (m. 16), *Rocking* (m. 120), and *Wild and Impassioned* (m. 206). Although much of the piece is structured upon a steady beat, I noticed that many of my favorite sections were written with complex changing meters. I felt the rhythms in Hyla’s music were a significant component and later learned that Hyla’s sense of rhythm was imbued with influence from Elliott Carter as well as nonclassical artists including James Brown (funk), Neil Young (singer/songwriter), Captain Beefheart (avant-garde rock), and Cecil Taylor (free jazz piano). These artists were outliers in their own genres, and rhythm is a salient feature of all of their music. Hyla’s complex rhythms felt grounded despite a lack of a steady underlying pulse. This temporal aesthetic is not unique to his music, but little research exists to meaningfully describe structure in music without isochronous meter.

I believed if I addressed how I was perceiving the music, then I would be able to phrase it in a way that felt natural to the listener. To me, the melodic lines sounded like an instrumental version of speech, and I began to explore how an understanding of linguistics could relate to the performance and perception of...
contemporary music without an isochronous beat. I began my investigation with Fred Lerdahl and Ray Jackendoff’s *Generative Theory of Tonal Music* (GTTM) because it is reflective of principles in universal psychological processes that inform musical analysis (Lerdahl & Jackendoff 1983). Although GTTM addresses the perception of rhythm and time through generative models borrowed from linguistics, the theory applies to tonal music and does not address much of the music being written today.

Without a meaningful rhythmic analysis from GTTM, I turned to linguistics, which explains the experience of rhythm according to perceptual salience because spoken syllables do not align with an even metrical grid and do not have even rhythms. Specifically, I am referring to metrical stress theory, which was first formalized in 1977 by Mark Liberman and Alan Prince, expanded in Morris Halle and Jean-Roger Vergnaud’s *An Essay on Stress* (1987), and further codified by Bruce Hayes in his 1995 book. Hayes’s book describes rhythm and emphasis in language as the culmination of his work in phonology regarding linguistic stress and metrical structure. I theorize that listeners experience rhythm in music without a steady pulse in the same way that they experience the rhythm of speech. Although stress theory was codified in English, the theory applies to many languages that organize according to stress; Hayes’s analyses include examples from numerous languages.¹ The structures elucidated by his theory do not apply only to English speakers.

My analyses outline structure and phrasing in music without periodic meter to aid performers.² I developed this theory through careful study of late twentieth-century music written in the United States of America, but I believe the result to be more widely applicable. My theory is an extension of GTTM with Hayes’s stress theory, and I present my final analyses in the form of tree structures that indicate the hierarchy of perceptual prominence within an excerpt.

### 1.1 General Structure

I must note three foci in discussing perception: the notated score, the performer’s mental construction, and the listener’s perception, the last of which is the focus of my analyses. The score is an intermediary device between the conception and the realization of a piece, so the notated meters do not necessarily relate to how the listener hears the music. This suggests the listener experiences two

---

¹ A representative list: Arabic, Cahuilla, Wargamay, Maithili, Hindi, Lenakel, Auea, Icelandic, Hixkaryana, Choctaw, Chickasaw, Creek, Seminole, Winnebago, Cayuga, Pacific Yupik, and Asheninca, among others.

² Unlike the repertoire discussed in Fred Lerdahl’s (1992) seminal work “Cognitive Constraints on Compositional Systems,” I discuss music that is cognitively accessible.
simultaneous structures: one created by the meter and another arising from the organization of prominent events on the musical surface. These two structures are often aligned in older repertoire, but temporal interest in metrical music is created through the misalignment of the meter and prominent events on the musical surface. However, periodic meter eclipses the prominence structure, which is only described in relation to the meter rather than recognized as an independent entity. For instance, syncopation and hemiola are two types of misalignment between these structures, but neither are disorienting for the listener because the prominent events are understood in relation to the regularity of the metrical structure. In music without a periodic meter, the listener must attend to the structure suggested by the events themselves rather than placing the rhythms in relation to an underlying meter; my analyses seek to foreground and describe this prominence structure. I do not seek to summon structures that were previously nonexistent, but rather to elucidate the prominence structure that is almost entirely overshadowed when periodic meter is present. In contemporary music without periodic meter, phrasing must thus rely entirely on the prominence structure.

To relate phrasing to temporal structures, my analytical approach extends GTTM with Hayes’s stress theory as a guiding principle, to focus on perceptual models that unite music and language. GTTM contains four components: grouping structure, metrical structure, time-span reduction, and prolongational reduction. However, I only address grouping and metrical structures; time-span and prolongational reductions rely on the same structures but focus on harmonic information, which is not the goal of my analyses.

Although my methodology involves linguistics, my analyses here do not include repertoire that involves speaking or singing. My focus is on deeper experiential connections between language and music, and works including conscious rhythmic connections between the two domains present a different analytical interest. In such pieces, the music is explicitly affected by language, whereas my work explores underlying connections between music and language.

Although my analyses rely on rhythm as a central factor, I address pitch and harmony tangentially as they relate to perceptual prominence. My analyses highlight perceptible structures related to phrasing by focusing on the salience of rhythm and meter as organizing principles. Although I involve elements of pitch in my analyses, I do not attempt to codify harmonic structure in the works from which I draw the musical examples. My theory applies to the phrase level of a piece and does not extend to higher hierarchical levels, such as the perceived salience of events within sections of an entire piece.
GTTM, the foundation of my theory, was co-authored by composer Lerdahl and linguist Jackendoff, who is also a clarinetist. The theory described perception of musical structures inspired by Noam Chomsky’s generative linguistic theory (Chomsky 1957, 1965, 1968). GTTM was grounded in the authors’ musical intuitions of heard structure in tonal music, and many of its principles have been successfully tested empirically. In order to describe analogous heard constructions in post-tonal music, I extend the theory with linguistic ideas associated with Hayes’s metrical stress theory.

GTTM assumes a regular periodic meter and a recurring beat hierarchy in each measure, so temporal tension must be described differently without these two conditions. Employing another linguistic theory to describe emphasis in post-tonal music aligns with the theory’s origins. As linguistic stress is hierarchical without evenness, the concept is useful when analyzing music without a constant, evenly spaced beat. Metrical stress theory suggests stress in language manifests as linguistic rhythm. Examining rhythm, meter, and temporality in a post-tonal context using widely accepted perspectives from linguistics contributes to an understanding of phrasing, in terms of both interpretation and perception, and provides a framework within which to comprehend structure.

Extending GTTM with constructs from language does not distort it, as the theory itself is deeply indebted to linguistic theory. My analyses are based on the points in time that draw the listener’s focus as based on emphasis in lieu of an underlying metrical structure as a grounding force. Events with greater perceptual salience are acknowledged as receiving greater emphasis within the prominence hierarchy. The emphasis that signals perceptual salience is more than just phenomenal accent, and I analyze music according to perceptual systems that structure stress in language.

Stress translates to music because the concept is not unique to language; it is a perceptual rather than physical phenomenon. Lerdahl mentions that stress is an important marker in both language and music because the perception of both syllables and pitch events “is one of relative sonic prominence within its immediate context” (Lerdahl 2001: 339). In both domains, no single physical correlate directly reflects stress or emphasis; the concept “unifies in a coherent way a broad set of phenomena” (Hayes 1995: 9). The perceptual markers of stress, in order of salience, are pitch contour, duration, and loudness, the last of which is relatively unimportant (Fry 1955, 1958). Specific characteristics of contour, duration, and loudness do not invariably signal stress; for example, high pitches do not always signal stress. In English, pitch is one of the most

---

powerful cues for stress, but the lowest pitch in the intonational contour of yes/no questions is the most stressed, whereas higher pitches are correlated with stress in other contexts. Lerdahl and Jackendoff define local stress as accent or “extra intensity on the attack of a pitch-event,” citing several forms of notated accents (Lerdahl & Jackendoff 1983: 78). Stress is not dependent on the specific linguistic structures of words and syllables because rhythms can be detected in many formats such as steps on pavement, machinery in motion, musical notes, or speech. Also, stress outlines a temporal structure that does not assume evenness or periodicity.

The primary goal of my research is to discuss temporal tension in music without an even beat, exploring how linguistic stress contributes to the analysis of recently written works. I focus on perceptible structures in order to relate the listener’s hearing of the music to practical performance applications.

2 On Prominence: Language and Music

Studies have shown a relationship between the rhythm of a composer’s language and their music by comparing the durational variability of each. For instance, a sequence of quarter notes with one eighth note will have very little durational variability as compared to a combination of notes of different durational values. Languages work similarly; each language is referred to as rhythmically organized around either stress or syllables. Stress-timed languages, such as English, German, Russian, and Arabic, organize around stresses, in that syllables vary greatly in length. For example, each word in the sentence “cats chase birds” includes a perceptual focus and relative rhythmic evenness that persists even when expanding the sentence with more words, such as “the cats are chasing the birds”; the added words are spoken with much shorter syllables. Syllable-timed languages, such as French, Italian, Finnish, and Turkish, are spoken with each syllable lasting roughly the same amount of time. Different amounts of durational variability can be measured with the normalized pairwise variability index (nPVI). This metric measures durational variability in sequences of attack points and has traditionally been applied to language. However, the nPVI can also measure durational variability in music.

Aniruddh Patel and Joseph Daniele showed a link between language and musical rhythm by examining works by English and French composers. English is stress-timed, and French is syllable-timed; this is reflected in the nPVI because English speech has a higher durational variability than French. This relationship is also reflected in English and French music, as shown in Ex. 1. Patel and Daniele drew source material from the Barlow and Morgenstern’s

---

Dictionary of Musical Themes (1983), choosing composers born in the nineteenth century who died in the twentieth century, native speakers of either British English or French who lived and worked in their respective countries. Each composer had to have at least five qualifying themes. Themes were excluded that referenced song or had external rhythmic or stylistic intentions, such as stylized dances or music replicating that of another culture. This comparison was successfully replicated with more composers and languages by David Huron and Joy Ollen, to show that the rhythmic variability of composers’ native languages affects their use of musical rhythm (Huron & Ollen 2003).

These qualities inform my examination of music with uneven rhythmic values and infrequent or varied repetition. Speech breaks down by group into sentences, phrases, and words, which can be understood as analogous to musical phrases and gestures. Because the rhythm of English speech is organized around stresses, syllables vary greatly in length. Curated speech, such as poetry or prepared speech, is often presented with a consciously controlled rhythm, but I am focusing on the rhythm of unplanned speaking.

Words in language and pitch in music can obscure the temporal similarities, but the speech-like quality of rhythms is perceptible through an aural comparison. To show the similarities, the following discussion juxtaposes examples of English speech and an excerpt from one of Hyla’s piano trios. I highlight the rhythm of the following examples by removing words and pitches, reducing each syllable and note to a single click. Because I have not located a recording of Hyla speaking, I am using my own speech as I am a native English speaker with a geographic trajectory similar to Hyla’s. Audio Example 1 contains audio of my reading aloud the first sentence of this paragraph. The sentence groups into small fragments, like gestures, and two main sections, like sub-phrases.

---

Example 1 Proportional relationship of English and French speech and music

Dictionary of Musical Themes (1983), choosing composers born in the nineteenth century who died in the twentieth century, native speakers of either British English or French who lived and worked in their respective countries. Each composer had to have at least five qualifying themes. Themes were excluded that referenced song or had external rhythmic or stylistic intentions, such as stylized dances or music replicating that of another culture. This comparison was successfully replicated with more composers and languages by David Huron and Joy Ollen, to show that the rhythmic variability of composers’ native languages affects their use of musical rhythm (Huron & Ollen 2003).

These qualities inform my examination of music with uneven rhythmic values and infrequent or varied repetition. Speech breaks down by group into sentences, phrases, and words, which can be understood as analogous to musical phrases and gestures. Because the rhythm of English speech is organized around stresses, syllables vary greatly in length. Curated speech, such as poetry or prepared speech, is often presented with a consciously controlled rhythm, but I am focusing on the rhythm of unplanned speaking.

Words in language and pitch in music can obscure the temporal similarities, but the speech-like quality of rhythms is perceptible through an aural comparison. To show the similarities, the following discussion juxtaposes examples of English speech and an excerpt from one of Hyla’s piano trios. I highlight the rhythm of the following examples by removing words and pitches, reducing each syllable and note to a single click. Because I have not located a recording of Hyla speaking, I am using my own speech as I am a native English speaker with a geographic trajectory similar to Hyla’s. Audio Example 1 contains audio of my reading aloud the first sentence of this paragraph. The sentence groups into small fragments, like gestures, and two main sections, like sub-phrases.

---

Example 1 Proportional relationship of English and French speech and music

Dictionary of Musical Themes (1983), choosing composers born in the nineteenth century who died in the twentieth century, native speakers of either British English or French who lived and worked in their respective countries. Each composer had to have at least five qualifying themes. Themes were excluded that referenced song or had external rhythmic or stylistic intentions, such as stylized dances or music replicating that of another culture. This comparison was successfully replicated with more composers and languages by David Huron and Joy Ollen, to show that the rhythmic variability of composers’ native languages affects their use of musical rhythm (Huron & Ollen 2003).

These qualities inform my examination of music with uneven rhythmic values and infrequent or varied repetition. Speech breaks down by group into sentences, phrases, and words, which can be understood as analogous to musical phrases and gestures. Because the rhythm of English speech is organized around stresses, syllables vary greatly in length. Curated speech, such as poetry or prepared speech, is often presented with a consciously controlled rhythm, but I am focusing on the rhythm of unplanned speaking.

Words in language and pitch in music can obscure the temporal similarities, but the speech-like quality of rhythms is perceptible through an aural comparison. To show the similarities, the following discussion juxtaposes examples of English speech and an excerpt from one of Hyla’s piano trios. I highlight the rhythm of the following examples by removing words and pitches, reducing each syllable and note to a single click. Because I have not located a recording of Hyla speaking, I am using my own speech as I am a native English speaker with a geographic trajectory similar to Hyla’s. Audio Example 1 contains audio of my reading aloud the first sentence of this paragraph. The sentence groups into small fragments, like gestures, and two main sections, like sub-phrases.

---

Example 1 Proportional relationship of English and French speech and music

Dictionary of Musical Themes (1983), choosing composers born in the nineteenth century who died in the twentieth century, native speakers of either British English or French who lived and worked in their respective countries. Each composer had to have at least five qualifying themes. Themes were excluded that referenced song or had external rhythmic or stylistic intentions, such as stylized dances or music replicating that of another culture. This comparison was successfully replicated with more composers and languages by David Huron and Joy Ollen, to show that the rhythmic variability of composers’ native languages affects their use of musical rhythm (Huron & Ollen 2003).

These qualities inform my examination of music with uneven rhythmic values and infrequent or varied repetition. Speech breaks down by group into sentences, phrases, and words, which can be understood as analogous to musical phrases and gestures. Because the rhythm of English speech is organized around stresses, syllables vary greatly in length. Curated speech, such as poetry or prepared speech, is often presented with a consciously controlled rhythm, but I am focusing on the rhythm of unplanned speaking.

Words in language and pitch in music can obscure the temporal similarities, but the speech-like quality of rhythms is perceptible through an aural comparison. To show the similarities, the following discussion juxtaposes examples of English speech and an excerpt from one of Hyla’s piano trios. I highlight the rhythm of the following examples by removing words and pitches, reducing each syllable and note to a single click. Because I have not located a recording of Hyla speaking, I am using my own speech as I am a native English speaker with a geographic trajectory similar to Hyla’s. Audio Example 1 contains audio of my reading aloud the first sentence of this paragraph. The sentence groups into small fragments, like gestures, and two main sections, like sub-phrases.
Theory of Prominence

Audio Example 1 Hayslett speech

The following excerpt from the middle of Hyla’s piano trio *Amnesia Redux* (2002) shows the similarity of Hyla’s rhythms to speech. The music in AEx. 2 can be heard as small gestures that relate more closely to speech than to tonal music, in terms of rhythm.

Audio Example 2 Hyla *Amnesia Redux* excerpt

This speech-like quality of rhythm is often disguised by pitch and harmony, so the following two audio examples present these excerpts reduced to rhythm alone. Changes in pitch are not salient to the listener’s segmentation of the musical surface, so temporal structure is an important factor in analyzing perception and affect (Lalitte et al. 2009).

The similarity between Hyla’s rhythms and speech rhythms is an important link to investigate because it is a factor that affects the aesthetic of the music, and the following audio examples show the similarity between the rhythms by removing everything except the rhythm. Each tick represents a syllable in AEx. 3 and the inception of a note in AEx. 4.

Audio Example 3 Hayslett speech reduced to rhythm

Audio Example 4 Hyla *Amnesia Redux* excerpt, reduced to rhythm

---

6 I created AEx. 3 and AEx. 4 manually using Garageband.
Through these examples, I have shown similarity between the rhythms in English speech and in Hyla’s music, which exemplifies a connection between rhythm in language and rhythm in music without a periodic beat. Jackendoff summarizes the connection between musical and linguistic rhythm, writing that “the similarities between musical rhythm and linguistic prosody are striking—although neither is reducible to the other. The two are related more or less like fingers and toes” (Jackendoff 2003: 80). Research has proven a subconscious link between temporality in composers’ music and their native languages, and this motivated the use of linguistic theory in the creation of my framework. Whether or not composers consciously use speech-like rhythms, linguistics provides a useful perspective in analyzing the resulting constructions.

**3 Rules for Analysis**

Rhythmic vocabulary is part and parcel of musical discussions, but I here define a few important terms used in my analyses. *Pulse* refers to a sequence of perceptually prominent points in time. The beats within the pulse are typically evenly spaced, but they do not have to be as long as they follow a pattern. Rhythms occur in relationship to the pulse, which is maintained internally by the performer and can be perceived by the listener. A pulse theoretically exists at all hierarchical levels, but I discuss the term in reference to the tactus, or the level of the temporal hierarchy to which listeners most often tap. If the pulse is “temporally regular,” listeners are able to anticipate future beats (London 2012: 9). In tonal music with an evenly spaced tactus, the pulse is often a shared perception between the listener and the performer; this evenness is a salient structural factor of the tonal tradition. The pulse could also be uneven as long as it exists within a recurring pattern, such as a beat pattern of 3+3+2 often heard in rock music. The hierarchical level of pulse to which the majority of listeners would tap their feet or nod their heads is referred to as the *tactus* level; this varies widely between pieces of music.

When the tactus level is not evenly spaced in time or does not contain a regular pattern, the concept of pulse diverges into separate experiences for the performer and listener. The performer maintains an internal sense of time mimicking an uneven pulse that matches the notated music, thereby sustaining tempo and placing rhythms; the beats maintained by the performer are in accordance with the metrical organization of the piece whether or not the beats are evenly spaced. The listener does not perceive pulse if the music lacks temporal regularity or a pattern; points of perceptual emphasis replace the listener’s sense of pulse in these situations. Analytically, prominence structure replaces metrical structure.
Meter is notated on the musical surface as measures. In tonal music, measures typically contain the same number of evenly spaced beats throughout one piece or movement. Meter is “the anticipatory schema that is the result of our inherent abilities to entrain to periodic stimuli in our environment” (London 2012: 12), analogous to a “yardstick whereby we locate the musical events of the piece in a grid of time-points” (Lester 1986: 118). This grouping of the pulse allows for hierarchical metrical structure, but perception of meter changes without periodicity just as with pulse. The performer maintains a sense of the notated meter, but the listener is likely unaware of its exact delineations. In much contemporary music, including the excerpts I will analyze, meter is a notational tool to provide a temporal grid and communicate local structure; it is not intended to be a perceptual phenomenon for the audience.

Isochronous meter is structured with beats at the tactus level that recur at evenly spaced points in time, usually in measures with the same recurring time signature. However, isochronous meter can exist between measures that can all be divided evenly by the tactus pulse, such as 3/4 and 4/4 that can be divided into quarter notes. In actual performance, meter is hardly ever isochronous due to slight deviations.

3.1 The Experienced Listener

In tonal music, temporal structure relies heavily on a perceptible meter. Through this device, listeners anticipate future beats and relate surface rhythms to an underlying regularity. However, in music lacking rhythmic evenness and meter as an experienced structure, the perceptually prominent beats in the heard structure must be analyzed without a sense of entrainment to meter; prominence emerges as an independent framework. To address temporal structure within these sections of music, my analyses outline underlying structures and aid with phrasing. Perceiving rhythmic and temporal structure in both music and language involves related processes.

Lerdahl and Jackendoff suggest that “a piece of music is a mentally constructed entity” and that the goal of music theory should be “to explicate this mentally produced organization” (Lerdahl & Jackendoff 1983: 2). The listener’s musical intuitions and perception create the heard structure, a phenomenon that requires experience with the musical idiom in order to organize the sounds meaningfully, according to GTTM. The theory addresses the listener’s final mental construction of a piece rather than real-time processing while listening. Rules assign analyses to pieces rather than leaving it to the analyst to fit the theoretical specifics to a piece, but the rules are meant to allow an understanding of the value of different readings of a work; no two listeners will hear a piece of music in the same way. Although Lerdahl and Jackendoff use the concept of an experienced listener, they acknowledge certain universal principles of musical
grammar. Ultimately, their assessment of the success of GTTM is contingent upon the information it reveals about a specific piece, how it addresses the nature of tonal music and music in general, and how it relates to cognitive theory in a broader sense (Lerdahl & Jackendoff 1983: 4).

With inspiration from Chomsky, the theory is “based in part on the goals and methodology, but not the substance, of generative linguistics” (Lerdahl & Jackendoff 1983: 45). Generative linguistic theory characterizes people’s linguistic knowledge while speaking a language they know. Humans can understand and construct an infinite number of unique sentences without ever hearing them before, which involves an understanding beyond knowledge acquired through direct instruction alone. Similarly, people can process and relate to music they have never heard before. Many other parallels have been made between music and language, but GTTM focuses on this cognitive capacity.

A GTTM analysis compares the coherence of different possible perceptions of a piece to determine which are the more or less preferred interpretations or the structures that the idealized “experienced listener” would more or less likely attribute to the music (Lerdahl & Jackendoff 1983: 9). The theory contains structures based on well-formedness rules (WFRs) that provide the conditions for all possible analyses and preference rules (PRs) that suggest which of the possible analyses more closely match the heard structure. The two aspects of GTTM pertinent to the development of my methodology are grouping structure and metrical structure, each of which has its own set of WFRs and PRs. However, the process is not an automated, straightforward assignment of an analysis because pieces invariably present conflicts between the rules. Broader context and relationships guide the prioritization of different rules in each situation. In this way, each analysis reflects some degree of interpretation, but my approach is also based on rules to describe structures within the music.

The primary component of GTTM that I employ is grouping structure, about which the authors write that “the rules for grouping seem to be idiom-independent—that is, a listener needs to know relatively little about a musical idiom in order to assign grouping structure to pieces in that idiom” (Lerdahl & Jackendoff 1983: 36). Irène Deliège proved that the rules for grouping structure apply broadly between musicians and nonmusicians alike (Deliège 1987: 356). Lerdahl and Jackendoff also recognize grouping structure as one of the most important factors that performers can manipulate to project their “largely unconscious) preferred analysis” of the pieces they perform (Lerdahl & Jackendoff 1983: 63–64). Because grouping allows for interpretive decisions by performers and is a structure perceptible to listeners without idiomatic knowledge of music, the concept is essential to laying out musical phrasing, and performers must be aware of how their articulation and breathing will affect the heard structure.