The Electrophysiological Correlates of Rhythm and Syntax in Music and Language

by

Harim Jung
Class of 2016

A thesis submitted to the faculty of Wesleyan University in partial fulfillment of the requirements for the Degree of Bachelor of Arts with Departmental Honors in Music and Psychology

Middletown, Connecticut April, 2016
Table of Contents

Abstract ........................................................................................................................................................................... 2

1 Introduction .................................................................................................................................................................. 3
  1.1 Overview .............................................................................................................................................................. 3
  1.2 Defining Syntax in Musical Harmony .................................................................................................................. 4
  1.3 Defining Rhythm in Music and Language ........................................................................................................ 7
  1.4 Overview of Auditory and Musical Event-Related Potentials (ERPs) .............................................................. 8
  1.5 Overview of the Shard Syntactic Integration Resource Hypothesis (SSIRH) ................................................. 12
  1.6 Overview of Dynamic Attending Theory (DAT) ............................................................................................... 14
  1.7 Integration of Theoretical Frameworks ........................................................................................................... 17
  1.8 Aims and Overall Predictions .......................................................................................................................... 18

2. Experiment 1 .......................................................................................................................................................... 20
  2.1 Materials & Methods ........................................................................................................................................ 20
    2.1.1 Participants .................................................................................................................................................... 20
    2.1.2 Materials and Stimuli ................................................................................................................................ 21
    2.1.3 Procedure ................................................................................................................................................... 23
    2.1.4 EEG Recording ......................................................................................................................................... 23
  2.2 Results ............................................................................................................................................................... 26
    2.2.1 Behavioral Results ..................................................................................................................................... 26
    2.2.2 ERP Results ............................................................................................................................................... 26
    2.2.3 Time Region Analyses ............................................................................................................................. 27
    2.2.4 Main Effects of Musical, Linguistic, and Rhythmic Violations ................................................................. 27
    2.2.5 Interactive Effects of Expectancy Violations ............................................................................................. 30
  2.3 Discussion .......................................................................................................................................................... 32

3. Experiment 2 .......................................................................................................................................................... 36
  3.1 Materials & Methods ........................................................................................................................................ 37
    3.1.1 Participants .................................................................................................................................................... 37
    3.1.2 Materials & Stimuli ................................................................................................................................ 38
    3.1.3 Procedure ................................................................................................................................................... 39
    3.1.4 EEG Recording ......................................................................................................................................... 39
  3.2 Results ............................................................................................................................................................... 40
    3.2.1 Behavioral Results ..................................................................................................................................... 40
    3.2.2 ERP Results ............................................................................................................................................... 40
    3.2.3 Time Region Analyses ............................................................................................................................. 41
    2.2.4 Main Effects of Musical, Linguistic, and Rhythmic Violations ................................................................. 41
  3.3 Discussion .......................................................................................................................................................... 43

4. General Discussion .................................................................................................................................................. 46

Acknowledgments ...................................................................................................................................................... 50

References .................................................................................................................................................................. 52
Abstract

Music and language are human cognitive and neural functions that have been shown to share neural resources in syntax processing (Patel, 2003) as well as temporal processing (Large & Jones, 1999). Although recent studies have investigated the sharing of neural resources for music and language, little is known about how music and language processing might interact as syntax unfolds over time. The current electroencephalography (EEG) study investigates the relationship between rhythmic expectancy and musical and linguistic syntax by presenting sentences, broken down into segments, paired with musical chords (adapted from Slevc et al., 2009). Linguistic syntax violations appeared in a garden-path design, and musical expectation violations, presented as out-of-key chords, and rhythmic expectancy violations, through early and late temporal perturbations, were manipulated at the critical region. Participants read sentence segments and listened to the musical chords, and answered questions about the sentences while their EEGs were recorded. Results show that musically irregular chords and linguistically unexpected garden-path sentences elicited an early anterior negativity (EAN), but tend to diverge over time, with a posterior positivity for musically irregular chords (P3a) and a late positive component (LPC) for linguistically unexpected sentences. Results also show that the N400 decreases in amplitude between pre-critical, critical, and post-critical time regions, suggesting re-analysis of semantic content throughout the course of reading sentence segments. Together, results suggest that the interaction of music and language syntax processing depends on rhythmic expectancy, which in turn affects attentional entrainment.
1 Introduction

1.1 Overview

Although music and language are both universal to human cognition, the degree to which these domains share cognitive resources is a long-standing debate in cognition and topic of discourse across and between several academic disciplines. Drawing from evidence from the fields of archeology, anthropology, and neuroscience, theorists have argued for music being a domain encoded into the human genome that shares evolutionary origins with language (Mithen, 2005). Such shared origins have been traced to extensive structural similarities between these two domains, as well as evidence for an evolutionary precursor to music and language (Botha, 2009). These structural similarities have specifically been investigated through the syntactic hierarchies by which discrete musical and linguistic elements are organized and combined. Just as linguistic syntax governs how words and phrases are arranged in language, syntactic structures in other domains, such as music, define the arrangement of elements in their respective domains. In music, such syntax could be understood as the rules that define how pitches are organized to form melody and harmony (Lerdahl & Jackendoff, 1983).

Despite these similarities, it has also been suggested that music and language have domain-specific components that challenge evolutionary models. By investigating neuropsychological evidence from populations with music-related deficits, theorists who support domain-specificity of music and language argue that music is a specialized function that is modularly processed (Peretz & Colteart, 2003). Given the domain-general vs. domain-specific debate between music and language
processing, there has been extensive research and evidence that examines the shared syntactic structures and processing mechanisms between music and language that rely on similar neurocognitive resources (i.e. Patel, 2003; Koelsch et al., 2005; Slevc et al., 2009). Moreover, because musical and linguistic structures unfold syntactically over time, the current study aims to investigate the shared neural resources in music and language syntax processing, and how such syntactic resources are affected by rhythm and temporal processing mechanisms.

1.2 Defining Syntax in Musical Harmony

Although syntax is commonly understood as the rules in which linguistic elements are hierarchically arranged to form sentences, it has been postulated that music has its own syntax as well. In the current study, musical syntax will be investigated by manipulating Western tonal harmony. In this system of harmony, pitches, which refer to the degrees in the perception of sound wave frequency, are organized into twelve chromatic semitones per octave. The octave is defined as the recurrence of a pitch class heard in a different frequency range based on a 2:1 ratio (Kostka et al., 2013). These chromatic pitches are organized into diatonic major and minor scales that the major and minor keys are built upon. These musical keys, which are organized relative to a central or “tonic” note, are not only the building blocks that govern the organization of pitches in Western tonal music, but also have been extensively psychometrically tested and mapped for the expectation and relationship of pitches within and between musical key areas (e.g. Krumhansl & Kessler, 1982; Lerdahl & Krumhansal, 2007). This suggests that this system of Western tonal harmony and the circle of fifths that defines the relationship between major and minor
key areas is learned syntactically for those who are exposed to Western music in a way similar to the learning of grammar in language.

Moreover, Western tonal harmony is defined by the arrangement of these pitches around a tonal center, or a key, in which such pitches are organized into melodic content, based on the sequential presentation of pitches in a horizontal temporal hierarchy, or harmonic content, based on the simultaneous presentation of pitches in a vertical temporal hierarchy (Kostka et al., 2013; Loui et al., 2005). This systematic organization of pitches based on the relationship, in terms of interval size and frequency of occurrence, between the tonic, or the pitch in which a musical key is centered around, and the other pitches within a key or tonal area, has thus been theorized and labeled as “tonality” (Dalhaus et al., 2009). The most fundamental of musical chords in Western music, or the triad, is composed of three notes that divide an interval of a 5th into two sets of thirds. Based on the quality of these superimposed thirds and the tonic note of a given triad’s relationship to the key it is in, the triad is defined as major, minor, augmented or diminished. These chords are then organized into chord progressions, in which certain chords occur more frequently to emphasize the key area that a given piece is in, and others are used less frequently to build drama or in pivotal sections, thus giving rise to a system of harmonic expectancy, in which some chords are “expected” and others are “unexpected.”

Among the most commonly occurring chords are the tonic chord, which is built upon the tonic note of the key a given passage of music is in and is represented by a roman numeral “I” in a major key, the dominant chord, which is built upon the fifth scale degree of the given key and is represented by a roman numeral “V,” and the
subdominant, which is built upon the fourth scale degree and is represented by a roman numeral “IV” (Kostka et al., 2013; Loui et al., 2005). Harmonic progressions typically follow a tonic-pre-dominant-dominant-tonic structure that serves a prototype for harmonic expectancy in musical syntax. Less common chords in a given major key, such as the ii, iii, vi, and vii\(^6\) chords, can serve as either substitutes for a subdominant or dominant function chord or have pre-predominant functions (see Figure 1).

![Diagram of typical chord progression structure in Western harmony](image)

Figure 1. Diagram of typical chord progression structure in Western harmony (adapted from Kostka et al., 2013)

In addition to the chords in a given key area, chords from other keys can be borrowed and also be used as substitutes for pre-dominant or dominant function chords, or in pivotal moments, such as during key changes or modulations. Examples of such chords include the Neapolitan sixth (N\(^6\)) chord, which is a major chord that is a semitone above the tonic chord (i.e. D-flat major chord in the key of C major) that serves as a predominant substitute, and the secondary dominant chord, which is the major V chord in the key of the dominant chord (i.e. D major in the key of C major) that serves as a dominant substitute during a modulation to the dominant key (Kostka et al., 2013). Thus harmonic expectation, or the rules of Western musical syntax, are violated when a chord, either in-key or out-of-key, is presented in a harmonic position where it is unexpected or not serving its proper function. In the current study, musical syntax violations are defined as out-of-key chords that are characterized by deviations
from harmonic function within the context of a chord progression in a major key. Such a violation of harmonic expectation has been extensively studied using behavioral and neuroimaging measures that clearly demonstrate a psychological detection of violation and used for comparing violations across different cognitive domains (i.e. Patel et al, 1998; Loui et al., 2005; Slevc et al., 2009, Koelsch et al., 2005).

1.3 Defining Rhythm in Music and Language

Rhythm describes the temporal aspects of auditory stimuli or music. In Western music theory, rhythm can be analyzed in terms of beat, tempo, and meter. The beat refers to the pulse of a given musical passage, and tempo refers to the rate at which a certain beat occurs. Such beats, moreover, are organized into recurring patterns or groupings known as meter based on patterns of stress in terms of accents and harmony. Meters can be defined by type based on the number of beats per measure – duple, triple, or quadruple – and as either simple or compound based on whether or not a beat can be divided into two or three equal parts (Kostka et al., 2013). Moreover, it has been suggested that rhythm is not simply just codified from the duration of notes or auditory events, but is cognitively and hierarchically processed and inferred as musical sequences unfold (Lounget-Higgins & Lee, 1982). Drawing from theories on rhythm in Western music the field of music cognition defines rhythm as the pattern of time intervals in a stimulus sequence, and is usually perceived as the time between event onsets (Grahn 2012). Such patterns induce the experience of metrically organized strongly or weakly accented pulses or beats, which are regular intervals that are not necessarily present in the auditory signal (Fitch,
2013). As a pattern of durations that engenders expectancies, rhythm may share commonalities with syntax and thus be processed similarly to both musical and linguistic syntax in the brain (Fitch, 2013; Patel, 2003). On the other hand, it has also been suggested that rhythm is an implicitly processed feature of environmental events that affects attention and entrainment to events in various domains such as music and language (Large & Jones, 1999). In order to understand how rhythm affects musical and linguistic syntax processing, the current study considers two major theories and tests their respective predictions using event-related potential (ERP) analysis technique in an EEG experimental paradigm.

1.4 Overview of Auditory and Musical Event-Related Potentials (ERPs)

Electroencephalography (EEG) is a neuroimaging technique with high temporal resolution that is sensitive to changes in electric potential by the use of electrodes distributed across the scalp. Such changes in potential are elicited by exogenous, or stimulus-driven, signals, endogenous, or task-dependent neural or motor signals, or the preparation and execution of a motor response to a given stimulus ( Luck, 2014). The average of these changes across the repeated presentation of a given stimulus results in the event-related potential (ERP). These ERPs, moreover, vary in amplitude, or the measure of activation synchrony of a given population of neurons, latency, or the time interval between the presentation of a stimulus and the peak of the component of interest, and duration, or the time between potential changes and return to baseline ( Luck, 2014). Because of the high temporal resolution of EEG neuroimaging technique, these ERPs range from short-latency components that reveal sensory-level aspects of signal processing, such as the Auditory Brainstem Response
(ABRs), to long latency components associated with higher-level cognitive structures and networks, such as the ERAN and the P600.

The current study investigates the role of expectancy violations on long-latency ERP components, which are highly dependent on factors of attention and expectation and give insight into how stimulus processing affects cognition. The N1 component, one of the earliest long-latency ERP response to auditory signals, is a negativity that peaks around 80-110ms, is selective to attention, and is associated with auditory perception (e.g. Hillyard et al., 1973, Luck 2014, Näätänen et al., 1981; Zobel et al., 2015). Moreover, the N1 component depends on the psychoacoustic properties of sound, such as pitch (Hillyard et al., 1973; Horváth, 2015; Näätänen et al., 1981), timbre (Moreno & Bidelman, 2014), and temporal expectancy (Besson & Faïta, 1995; Escoffier et al., 2015), and is modulated by deviants in chord sequences (Virtala et al., 2014), various properties of speech such as emotional prosody, phonetics, and vowel perception (Bidelman & Alain, 2015; Pinheiro et al., 2015; Polat & Atas, 2014), as well as properties of auditory stream segregation (Francois et al., 2014; Zobel et al., 2015). The auditory N1 component is then followed by a positivity (P2), peaking in latency at 150-275ms after event onset, that is associated with auditory learning, and is enhanced by deviant tones in auditory oddball tasks (e.g. Besson & Faïta, 1995; Hillyard et al., 1973; Moreno & Bidelman, 2014, Näätänen et al., 1981). The N1 and P2 components together form an N1-P2 complex that is modulated by delays in rhythmic expectancy, and is dependent on attention in amplitude (Besson & Faïta, 1995; Hillyard et al., 1973).
This complex is followed by a secondary negativity (N2) that peaks 160-220ms after event onset and is elicited by a deviant tone in a stream of repetitive non-target stimuli (Luck, 2014; Näätänen et al., 1981). One subcomponent of this ERP, the N2a, also known as the mismatch negativity (MMN), is elicited to auditory mismatches in intensity, temporal expectancy, omission of a tone, or deviant tones (Luck, 2014). Unlike the N1-P2 complex, the MMN is independent of attention and is elicited in a task-independent, unconscious, pre-attentive matter, as seen in dichotic listening tasks (Näätänen et al., 1981) and tests with distractions in rhythmic structure or loudness of tone sequences (Bouwer et al., 2014; Miller et al., 2015). However, it has been shown that the anteriorly distributed and endogenous N2 component (N2b) is in fact modulated in latency by attention – specifically, a faster latency has been demonstrated for attended deviant tones (Näätänen et al., 1981). Moreover, it has recently been shown that MMN components elicited by attended auditory stimuli is followed by an intensified P3 deflection, and that attention affects the amplitude of polarity reversal in the frontocentrally distributed MMN to supratemporal and mastoidal regions (Miller et al., 2015).

In addition to the N2 component, deviant tones in elicit a late positivity (P3) that is also seen in the processing of musical expectancy violations (e.g. Besson & Faïta, 1995; Hillyard et al., 1973; Miller et al., 2015; Näätänen et al., 1981; Putkinen et al., 2014b). The P3 component, which peaks at 250 – 400ms after event onset, show different distribution patterns depending on attention, with the automatic and task-irrelevant P3a maximally distributed frontally, and the actively attended and task-relevant P3b maximally distributed parietally (Hillyard et al., 1973; Luck, 2014).
Unlike the N2, the P3b amplitude and latency is dependent on the magnitude of the deviant tone violation (Hillyard et al., 1973; Näätänen et al., 1981). Similar late positive components (LPCs) have been found in studies using more complex musical stimuli, and was also seen when the final note of these musical excerpts were delayed by 600ms, suggesting that the late positivity (P3/LPC) is elicited by different expectancies that may share neurocognitive resources (Besson & Faïta, 1995).

Furthermore, ERP components associated with the processing of higher-level musical structures, such as harmony, show similar patterns of expectancy (Koelsch et al., 2000; Patel et al., 1998). The development of such expectations can be seen in a late negativity peaking at 500-5500ms after event onset (N5) that spikes in amplitude when musical expectancy is violated by an out-of-key chord, an early negativity peaking around 150ms that is maximal in anterior regions and is right-lateralized (ERAN), both of which are not dependent on attention and elicited by task-irrelevant harmonic violations (e.g. Patel et al., 1998; Koelsch et al., 2000). Moreover, it has been shown that this ERAN and a late positivity (P600), elicited by harmonic expectancy violations, have similar counterparts in the processing of linguistic expectancy violations (Friederici, 2002; Patel et al., 1998; Steinbeis & Koelsch, 2008). Given these similarities, as well as the effects of rhythm and attention on ERP components associated with lower-level auditory signal processing, the current study considers past research and theoretical frameworks that address the extent to which music and language rely on similar processing mechanisms and resources and how rhythm may share or tax similar resources based on attention.
1.5 Overview of the Shard Syntactic Integration Resource Hypothesis (SSIRH)

The Shared Syntactic Integration Resource Hypothesis (SSIRH), which postulates an “overlap in the neural areas and operations which provide the resources for syntactic integration,” has greatly influenced research in music and language (Patel, 2003). Drawing from and reconciling contrasting findings in neuropsychology and neuroimaging studies on syntax processing, the SSIRH suggests that the same syntactic processing mechanisms act on both linguistic and musical syntax representations. Moreover, the hypothesis predict that these syntactic processing resources are limited, and thus studies that integrate musical and linguistic syntax processing will show patterns of neural interference (Patel, 2003a). While there is ongoing debate concerning the nature of these shared resources (Slevc & Okada, 2014) and the extent to which these shared resources are specific to syntax (Perruchet & Poulain-Charronnat, 2012), convergent studies do provide evidence for some shared processing of music and language, with evidence ranging from behavioral manipulations of syntactic expectancy violations in music and language (e.g. Slevc et al., 2009; Fedorenko et al., 2009; Hoch et al., 2011) to neuroimaging methods such as ERP and EEG studies that track the neural processing of syntax and its violations (e.g. Koelsch et al., 2005; Steinbeis & Koelsch., 2008; Fitzroy & Sanders, 2013).

Behavioral research supporting the SSIRH suggest domain-generality for the processing of syntax in music and language, and that such neural resources are limited and specific to syntax. Such domain-general and syntax-specific resources are supported by reaction time studies that demonstrate increased reading times and interactive effects during simultaneous processing of musical and linguistic syntax.
expectancy violations, but not for musical syntax violations paired with linguistic semantic or timbral expectancy violations (Fedorenko et al., 2009; Slevc et al., 2009). Yet, contrasting behavioral evidence challenges the syntax-specific nature of these domain general resources by demonstrating similar effects for semantic garden-path sentences and interactions between linguistic and musical violations, regardless of linguistic violation type (Hoch et al., 2011; Perruchet & Poulin-Charronnat, 2013). Despite these findings, follow-up evidence aimed at codifying and testing the intersections of musical and linguistic semantics reveal that musical semantics do not effect linguistic semantics processing, and that semantics in these domains are incompatible in specificity, compositionality, and communication (Kolesch et al., 2011; Slevc & Patel, 2011).

Moreover, neuroimaging studies further support the domain general and limited resources hypothesis of the SSIRH by investigating linguistic and ERP components that demonstrate patterns of interaction between music and language syntax processing. ERP studies demonstrate that both musical and linguistic syntax violations elicit early anterior negativities (EAN), localized to the inferior frontal cortex, and a late positivity (P600), pointing towards these shared resources in music and language syntax processing (Patel et al., 1998; Hahne & Frederici, 1999; Friederici, 2002; Koelsch et al., 2000, Koelsch et al., 2005). The domain-general and syntax-specific resource hypothesis are further supported by a reduced ELAN during simultaneous processing of musical and linguistic syntax violations, and lack interactive effects on the semantic N400 in the simultaneous processing of linguistic semantic and musical syntax violations, as well as on the MMN in the simultaneous
processing of linguistic syntax violations and deviant tones (Koelsch et al., 2005). Additional neuroimaging evidence from specific populations extend these findings by demonstrating that musicians process syntax in a specialized manner in motor cortices (Fitrozy & Sanders, 2013; Jentschke & Kolesch, 2009; Sammler et al., 2009) and that children with specific language impairments (SLI) have deficits in musical syntax processing (Jentschke et al., 2008).

The current study aims to investigate the effect of rhythm on musical and linguistic syntax processing through the manipulation of rhythmic expectancy through early, on-time, or late conditions. In this ERP paradigm, we will consider the possibility of rhythm, a feature of music and stimuli presentation that engenders expectancy, having a syntactic organizational structure (i.e. Fitch, 2013). It has been shown that rhythm effects and enhances various aspects of musical and linguistic stimuli processing. These studies range from neuroimaging studies that show rhythmic structures enhance brain activity in areas associated with language in musicians (Herdener et al., 2014), as well as studies that show rhythmic priming enhances musical and linguistic syntax processing (Kentner, 2012; Przybylski et al., 2013). These results, as well as the current study, open up a door to investigate the possibility of rhythm as an influencer of musical and linguistic syntax processing. In order to investigate this, the development of rhythmic expectancies through attention and entrainment must be considered as well.

1.6 Overview of Dynamic Attending Theory (DAT)

As both musical and linguistic syntax unfold over time, the timing of both musical and linguistic events may affect such sharing of their processing resources. While
rhythmic structures develop expectations and have been suggested to be organized and processed similarly to musical and linguistic syntax, it has also been suggested that rhythm is an implicitly processed feature of auditory and visual events that affect attention and entrainment to such events (Large & Jones, 1999). The Dynamic Attending Theory (DAT) posits a mechanism by which internal neural oscillations, or attending rhythms, synchronize to external rhythms of the stimuli in our environments (Large & Jones, 1999). The entrainment model, moreover, suggests that rhythmic processing is a fluid process in which attention is involuntarily entrained in a periodic manner, to a dynamically oscillating array of external rhythms, with attention peaking with stimuli that align periodically with the oscillator (Grahn, 2012; Large & Jones, 1999). This process of attentional entrainment has been suggested to occur through neural resonance, or the process by which neurons form a circuit that is periodically aligned with the stimuli, which allows for the hierarchical organization of stimuli with multiple neural circuits resonating at different levels, or subdivisions, of the meter (Henry, Herrmann & Obleser, 2015; Large & Snyder, 2009).

Research supporting the dynamic attending theory has demonstrated that entrainment occurs across different perceptual modalities. In the auditory modality, comparative pitch judgement tasks in which interleaving tones separated temporally by regular inter-onset intervals (IOIs) were presented demonstrated attentional entrainment through reduced accuracy for distractor tones (Jones et al., 2002). More specifically, pitch judgments were found to be more accurate when the tone to be judged was separated rhythmically from the interleaving tones by a predictable IOI, than compared to those of early or late tones that were separated by a shorter or
longer IOI respectively (Jones et al., 2002). Such entrainment in the auditory modality occurs via anticipatory attending, or the process by which regular IOIs focus attention on anticipated points that produce peaks of internal attentional oscillations as these specific times (Jones et al., 2006). Likewise, attentional entrainment has been shown to occur in the visual modality, in studies that demonstrate enhanced rhythmic discrimination ability for rhythmically regular visual and auditory stimuli (De Freitas et al., 2014).

Furthermore, this entrainment model has been extended to the processing of higher level musical and linguistic structures. Studies on artificial grammar learning, for example, demonstrate that unit length regularity enhanced the learning of linguistic verbal and auditory nonverbal artificial grammars (Hoch et al., 2013). It has also been demonstrated that syntactic priming led to the development of perceptual expectations similar to that of rhythmic entrainment in tasks measuring implicit auditory sequence learning (Escoffier & Tillman, 2008; Tillmann & Poulin-Charonnat, 2010). This model of attentional entrainment is further supported by neuroimaging studies that investigate the effects of entrainment on higher-level music and linguistic syntax processing. Specifically, it has been shown that rhythmically irregular presentation of linguistic stimuli led to reduction in the P600 in linguistic syntax processing, a late component that is associated with syntactic integration and re-analysis and is dependent on attention (Schmidt-Kassow & Kotz, 2008). This suggests that the effect of rhythm on syntax processing may be due to attentional entrainment, and that rhythmic structures are processed earlier than their linguistic and musical syntax counterparts. Additionally, it has been shown that attention to
harmonic violations enhanced the amplitude and latencies of indices of musical syntax processing in ERP studies (Loui et al., 2005), suggesting implicit and automatic rhythmic entrainment is an important aspect syntax processing that must be further investigated (Loui et al., 2005; Henry, Hermann & Obleser, 2015).

### 1.7 Integration of Theoretical Frameworks

Both SSIRH and DAT make predictions about how our cognitive system processes events as they unfold within a stimulus sequence. While the predictions from the SSIRH pertain to expectations for linguistic and musical structure, those from the DAT pertain to expectations for temporal structure in streams of auditory and visual stimuli. Though these theoretical frameworks address fundamentally different domains of cognition, there has been growing interest in investigating the intimacy between rhythm and music and language processing. Such interest has yielded the question and ongoing debate on the extent to which rhythm is structurally alike and processed similarly to music and language syntax (Fitch, 2013; Patel, 2003), or is more an inherent and implicitly processed feature of audiovisual stimuli that affects attentional entrainment (Large & Jones, 1999). Therefore, if rhythm were processed similarly to musical and linguistic syntax, then a violation in rhythmic expectancy would elicit similar behavioral and neurological patterns of error processing, and patterns of interference in processing simultaneous violations in rhythmic and syntactic expectancies (Jung et al., 2015). On the other hand, if rhythm was more of an implicit aspect of music and language, then violations in rhythmic expectancy would, instead, yield temporal modulations to syntactic error processing.
Studies that examine the effects of rhythm on music and language processing challenge such a dichotomous view on rhythm, and call for an integration, rather than a separation, of these theoretical frameworks. Such studies include behavioral and ERP studies that demonstrate delayed behavioral responses and enhanced attenuation of ERP responses associated with syntactic integration (i.e. P600) for rhythmically unexpected musical and linguistic events (Schmidt-Kassow & Kotz, 2008; Bouwer & Honing, 2015; Jung et al., 2015), as well as studies that demonstrate enhanced syntactic error detection and processing with rhythmic priming and regularity of stimuli (Gordon et al., 2015). Moreover, the interconnection between rhythm and music and language is further demonstrated by developmental and clinical studies that demonstrate that rhythmic discrimination ability and rhythmic priming enhance the processing and detection of errors in musical and linguistic syntax, and that specialized populations, such as jazz musicians, recruit language-specific areas (increased activation of left SMG) for processing rhythmic structure (i.e. Gordon et al., 2015a; Kentner, 2012; Przybylski et al., 2013; Herdener et al., 2014). Because it has been clearly demonstrated that music and language syntax processing depends on rhythmic structure, the current study aims to further examine this effect of rhythm, and whether rhythmic expectancies are built upon syntactic structure and/or attentional entrainment.

1.8 Aims and Overall Predictions

The current study aims to examine the simultaneous cognitive processing of musical syntax, linguistic syntax, and rhythmic expectancies. We extend the reading time paradigm of Slevc et al (2009), by borrowing from the rhythmic expectancy
manipulations of Jones et al (2002), to investigate how the introduction of rhythmic expectancy affects musical and linguistic syntax processing. Rhythmic expectancy was manipulated through rhythmically early, on-time, or late conditions relative to a fixed, expected onset time. The effect of rhythmic expectancy on musical and linguistic syntax processing will be investigated through two experiments. The first experiment will investigate such effects in a paradigm in which manipulations in each of the three expectancies can occur either independently of one another, or simultaneously, in order to investigate the shared resources between music and language as hypothesized by the SSISRH and examine how rhythm taxes such resources through syntactic or attentional means. In the second experiment, we examine expectancy violations in musical syntax, linguistic syntax, and rhythmic expectancy independently in a counterbalanced design to not only investigate the individual effects violations of each factor in an ERP paradigm, but to have as control data to compare with the processing of simultaneous violations in Experiment 1.

In both Experiment 1 and Experiment 2, we expect to see musical and linguistic syntax violations to elicit their corresponding ERP components associated with syntactic error processing (EAN), and cognitive attentional re-orientation (P3a) and memory-related context updated (LPC/P600) following syntactically incongruent events (e.g. Patel et al., 1998; Koelsch et al., 2000; Hillyard et al., 1973; Besson & Faïta, 1995; Friederici, 2002; Steinbeis & Koelsch, 2008). As demonstrated by past research, we expect to see that violations of rhythmic expectancy through perturbations from temporal regularity will affect syntax processing and modulate the ERP components associated with syntax processing (Schmidt-Kassow & Kotz, 2008).
Moreover, we aim to compare the independent effects of rhythmic violations and compare them with the ERP components elicited by syntactic expectancy violations in music and language, as well as the effects of rhythmic violations on processing violations in music and language syntax. The current EEG study will thus allow us to more specifically examine how rhythmic expectancy may differentially modulate the processing of musical and linguistic syntax.

2. Experiment 1

2.1 Materials & Methods

Participants read sentences that were broken down into segments, each of which was paired with a chord from a harmonic chord progression. Linguistic syntax expectancy was manipulated using syntactic garden-path sentences, musical expectancy was manipulated using chords that were either in key or out of key, and rhythmic expectancy was manipulated by presenting critical region segments early, on time, or late.

2.1.1 Participants

30 undergraduate students (mean age (19.47 years), age range (18 – 21 years), SD = 1.25) from Wesleyan University participated in this study in return for course credit or monetary compensation. A recording error resulted in the loss of data for 4 out of the 30 total subjects and one out of the 26 remaining were excluded due to physiological artifacts, and so data for 25 participants were used in the final analysis. Of the remaining participants, all reported normal hearing 15 participants (60 %) reported having prior musical training, averaging 6.97 years (SD = 3.54), 13 (52%)
participants identified as female, and 12 as male. 22 (88%) reported that their first language was English, 6 were native speakers of English and one other language, and 3 had a language other than English as their first language. Other than English, the first languages of participants include Chinese and Korean. 9 participants (36%) spoke more than one language. All participants had normal or corrected-to-normal vision and reported no history of psychiatric or neurological disorders. Informed consent was obtained from all subjects as approved by the Ethics Board of Psychology at Wesleyan University. Subjects were excluded from the analysis if more than of their trials contained physiological artifacts (i.e. eye blinks, erratic eye movements, excessive muscle activity), or if more than half of their answers to reading comprehension questions were incorrect. As a result, one of the 26 subjects were excluded, leaving 25 in the final analysis.

2.1.2 Materials and Stimuli

Participants read sentences, from Slevc et al. (2009), which were broken down into segments of one or several words, each of which was paired with a chord from a harmonic chord progression. All stimuli were generated and presented using MaxMSP software (Zicarelli, 1998) using a MacBook Pro. Linguistic syntax expectancy was manipulated through syntactic garden-path sentences, musical syntax expectancy was manipulated through chords that were either in key or out of key, and rhythmic expectancy was manipulated thorough early, on time, or late presentations of critical region segments. Sentence segments and paired harmonic progressions were presented at the regular inter-onset interval (IOI) of 1200 ms with critical region segments presented either on time, or perturbed to be either early late. The early jitter
was 115 ms earlier than the on-time presentation, and the late jitter was 115 ms later. Therefore, the IOIs were either $1200 - 115 = 1085$ ms (early), $1200$ ms (on time), or $1200 + 115 = 1315$ ms (late) (Figure 2). Chord progressions, also from Slevc et al (2009), were played in MIDI using a grand piano timbre, followed the rules of Western tonal harmony, and were all in the key of C Major. Out-of-key chords violated harmonic expectancy given the context, but were not dissonant chords by themselves (i.e. N6). A yes-or-no comprehension question was presented at the end of each sentence. 96 unique comprehension questions, two for each sentence, were written so each sentence would have one comprehension question written to have a correct answer “yes,” and another to have a correct answer “no.”

![Schematic illustration of experimental design and stimuli presented in one trial.](image)

Figure 2. Schematic illustration of experimental design and stimuli presented in one trial.

12 unique experimental modules were created in order to counterbalance the experimental design. Each module contained all 48 sentences, with violation and filler conditions rotated through the sentences in order to control for effects of content, length, and sentence order. Each module contained: 4 rhythmic violation trials (2 early and 2 late), 3 musical syntax violations trials, 1 linguistic syntax violation trial,
5 musical syntax plus rhythmic violation trials, 1 linguistic plus musical syntax violation trial, 2 linguistic syntax plus rhythmic violation trial, 2 trials with all 3 violations, and 30 sentences with no violations. Therefore, in a given module, only 37.5% of trials contained any violation. Half of the sentences in a module were assigned a “yes” question, and the other half was assigned a “no.” Each subject completed 3 experimental modules (144 trials per subject) and the order of trials and experimental modules assigned were randomized for each subject.

2.1.3 Procedure

Before beginning the experiment, the participants gave informed consent and completed a short background survey. Participants were instructed to pay close attention to the sentences rather than the chord progressions, making the musical stimuli task-irrelevant. Then the participants ran through a set of practice trials. After the practice trials, in the actual experiment, the experimenter selected one of the 12 possible experimental modules at random. The following segments appeared at a fixed IOI regardless of when the current segment disappeared. After the end of each sentence, a yes-or-no comprehension question was displayed, at which point participants answered the question by pressing Y or N on the keyboard, ensuring participants were attending to the reading. Answering the comprehension question cued the next trial. The same procedure was repeated for the other two blocks. The experiment lasted about 60 minutes.

2.1.4 EEG Recording

EEG was recorded from 64-channel silver chloride electrodes mounted to a custom-designed elastic cap (Brain Vision LLC) and referenced to the FPz channel.
during recording. Electrode placement matched the International 10-20 system locations. Electrode impedances were kept less than or equal to 5kΩ across all channels. All channels were continuously recorded and digitized with a sampling rate of 1000 Hz. Recordings took place in an electrically shielded, sound-attenuated chamber, and the experiment was conducted with the lights off.

2.1.5 Data Analysis

Behavioral comprehension question response were saved as text files from MaxMSP, and imported into Microsoft Excel and SPSS for statistical analysis. ERPs were averaged separately for expected and unexpected stimuli, over a time window of 200 ms prestimulus to 1000 ms poststimulus. Infinite impulse response (IIR) filters were applied with a low cutoff of 0.5, a high cutoff of 30 Hz, and a notch filter of 60 Hz. Raw data inspection was used to exclude data points with a higher gradient (> 50uV/ms), higher maximum difference (> 200uV/200 ms), and extreme amplitudes (beyond -200 to 200 uV). The raw EEG data was also corrected for eye blinks and vertical eye movement using ocular correction independent component analysis in BrainVision Analyzer 2.0. The data was then segmented into different musical, linguistic, and rhythmic conditions, and the trials were averaged and baseline corrected. Difference waves between expected and unexpected trials were also taken for each subject. ERPs were grand averaged across the 25 acceptable subjects, and mean amplitudes were exported from BVA and imported into Excel and SPSS. Statistical analyses across subjects were performed on mean amplitude measures across specific latency windows.
Drawing from analysis techniques of past ERP studies (see Miranda & Ullman, 2007), ERPs were statistically evaluated over six regions of interest (ROIs; Figure 3): left anterior (F1, F3, F5, F7, AF3, AF7), midline anterior (AFz, Fz, FCz), right anterior (F2, F4, F6, F8, AF4, AF8), left posterior (P1, P3, P5, P7, PO5, PO7, O1), midline posterior (CPz, Pz, POz), and right posterior (P2, P4, P6, P8, PO4, PO8, OZ). Mean amplitudes for each condition were computed for three separate time windows for ERP components of interest. Time windows were selected based on previous research on musical and linguistic syntax and visual analysis of grand averaged data across all subjects showing components in question. Two-tailed paired-sample t-test and ANOVAs were thus calculated for each time window and ROI combination with the dependent variable of ERP component amplitude, fixed factors of music (in key vs. out of key), language (congruent vs. incongruent), and rhythm conditions (early vs. on time vs. late), and a random factor of subject number.

Figure 3. Montage of Regions of Interest (ROIs) used for statistical analyses of ERP data
2.2 Results

2.2.1 Behavioral Results

On comprehension questions, participants performed significantly above chance in all conditions (overall $M = 89.57\%$, $s = 4.49$, two-tailed t-test against chance level of 50% correct: $t(24) = 43.22$, $p < .0001$). Because none of the participants performed below 50% on the comprehension questions, all participants successfully attended closely to the linguistic stimuli, while ignoring the auditory stimuli.

2.2.2 ERP Results

Visual analysis of the grand averaged waveforms indicated auditory N1-P2 complexes for all conditions in both violated and non-violated conditions, as well as four ERP components elicited by the violation condition as opposed to the non-violated conditions. The time windows for these components selected based on previous research and include: 180-210ms for an early anterior negativity (EAN) elicited by out-of-key and linguistically incongruent trials, 450-500ms for a posterior positivity (P3a) elicited by out-of-key conditions, 450-500ms for a negativity (N400) that showed attenuation across time (pre-critical, critical, and post-critical segments), and 750-800ms for a posterior positivity (LPC) elicited by both out-of-key and linguistically incongruent trials (Patel et al., 1998; Koelsch et al., 2000; Hillyard et al., 1973; Besson & Faïta, 1995; Friederici, 2002; Steinbeis & Koelsch, 2008).
2.2.3 Time Region Analyses

Analyses were first performed across the pre-critical, and post-critical time regions in order to investigate possible effects of habituation due to repeated stimulus onset. These analyses were plotted on midline anterior channel FCz, and indicate a reduction of the P2 (180-210ms range) and N400 (450-500ms range) components across time (Figure 4). Two-tailed paired samples t-tests were conducted between pre-critical and post-critical segments for these components on channel FCz, confirming P2 ($t(24) = 5.39, p < .0001$) and N400 reductions ($t(24) = 6.19, p < .0001$). These reductions indicate attenuation or habituation late in each trial after critical musical and linguistic segments have occurred, suggesting reduced perceptual and cognitive resource allocation after the critical region of each sentence.

![Figure 4. ERP component analyses across pre-critical, critical, and post-critical time segments on FCz.](image)

2.2.4 Main Effects of Musical, Linguistic, and Rhythmic Violations

Visual inspection of the musically in-key vs. out-of-key waveforms at midline anterior channel FCz suggest that, compared to their in-key counterparts, musically out-of-key chords at the critical region elicited and early anterior negativity (EAN) at the 180-210ms time window, a middle positivity (P3a) at the 450-500ms time
window, as well as a late posterior positivity (LPC) at the 750-800ms time window (Figure 5). Two-tailed paired samples t-tests and repeated-measure ANOVAs were conducted across the established time windows and ROIs. Although there were no significant differences in average ERP amplitude or main effect of musical condition at the 180-210ms window, analyses indicate effects the musically out-of-key chords at later time windows. On average, the amplitudes of ERPs to out-of-key chords ($M = 0.72, SD = 3.54$) were more positive than those of their in-key counterparts ($M = -0.91, SD = 2.08$) at the 450-500ms time window at midline posterior electrodes ($t(24) = 2.56, p < .05$).

![Figure 5](image.png)

*Figure 5. ERP component analyses for musically in-key vs. out-of-key trials on FCz.*

One-way ANOVAs with the dependent variable of ERP amplitude, independent variable of musical condition, and a random factor of subject number confirm the P3a at midline anterior ($F(1,24) = 8.91, MSE = 83.85, p = .006$) as well as midline posterior ($F(1, 24) = 6.54, MSE = 123.83, p = .017$; see Figure 4). ANOVA tests also indicate a late posterior positivity (LPC) at the 750-800ms time window at the left posterior ($F(1,24) = 6.55, MSE = 75.36, p = .017$), midline posterior ($F(1,24) = 4.63, p = .017$).
\(MSE = 106.52, p = .042\), and right posterior \((F(1,24) = 8.14, MSE = 92.64, p = .009)\) ROIs.

Visual inspection of the linguistically congruent vs. incongruent waveforms at channels Fz and CPz suggest that, compared to their congruent counterparts, linguistically congruent sentence segments at the critical region elicited and early anterior negativity (EAN) at the 180-210ms time window and a late posterior positivity (LPC) at the 750-800ms time window (Figure 6). Although there were no significant differences in average ERP amplitude or main effect of linguistic syntax condition at the 180-210ms window, analyses suggest effects of linguistic syntax incongruity at the later time window. On average, the amplitudes of ERPs of linguistically incongruent sentence segments \((M = 0.85, SD = 2.52)\) were marginally more positive than those of their congruent counterparts \((M = 0.22, SD = 1.17)\) at the 750-800ms time window at left posterior electrodes \((t(24) = 1.76, p =0.09)\). Repeated-measure ANOVA tests with the dependent factor of ERP amplitude, independent factor musical condition, and random factor of subject number confirm this marginally significant LPC at the left anterior \((F(1,24) = 3.09, MSE = 79.76, p = .092)\) ROI.

![Figure 6. ERP component analyses for linguistically congruent vs. incongruent trials on Fz and CPz.](image)
Visual inspection of the rhythmically early vs. on-time vs. late waveforms at midline anterior channel FCz suggest that, compared to their on-time counterparts, rhythmically early and late sentence segments elicit an enhanced anterior positivity (P2) at the 180-210ms time window, and an enhanced P3a at the 450-500ms time window (Figure 7). Repeated measure ANOVA tests reveal that the effect of rhythmic expectancy violations on the P3a was only marginally significant on midline anterior channels ($F(2,48) = 2.91, MSE = 201.57, p = .064$). This suggests that rhythmic violations share some similarities with musical syntax violations in eliciting a P3a.

**Figure 7.** ERP component analyses for rhythmically early vs. on-time vs. late trials on FCz.

### 2.2.5 Interactive Effects of Expectancy Violations

Given that one-factor analyses for main effects of musical syntax, linguistic syntax, and rhythmic expectancy have only yielded significant effects of musical syntax expectancy on ERP amplitude for specific ROIs at specific time windows, two-factor and three-factor analyses were conducted to investigate the possibility of interactive effects at the critical region. Two-way ANOVAs with fixed factors of
language and music, music and rhythm, or language and rhythm, and random factor of subject number were used to test the interaction between pairs of the different expectancies at each of the time windows and ROIs. Results indicate that at the 450-500ms time window, there was an interaction between music and rhythm ($F(2,48) = 3.57, MSE = 1728.60, p = 0.036$) at the left posterior ROI. In order to further investigate this interaction, we plotted mean ERP amplitudes across rhythmically early, on-time, and late trials for musically in key and out of key conditions at the 450-500ms time window for left posterior channels (See Figure 8a). This analysis indicates that on average, mean amplitude was greater for out of key ($M = 5.36, SD = 7.96$), than in key ($M = 0.26, SD = 8.52$) trials ($t(24) = 2.85, p = 0.0087$) for rhythmically early trials. Moreover, as indicated by the plot, both early and late rhythmic violations enhance the amplitude of the P3a in musically out-of-key trials, while reducing it for musically in-key trials (Figure 8).

Figure 8. **Rhythmic effects on music & language**: Mean ERP amplitudes for musically in key and out-of-key trials (a) and linguistically congruent and incongruent trials (b) during early, on time, and late conditions. Error bars show standard error.

Results also indicate an interaction between language and rhythm at the 450-500ms window ($F(2,48 = 4.25, MSE = 1282.75, p = 0.02$) at the left anterior ROI. In
order to further investigate this interaction, we plotted mean ERP amplitudes across rhythmically early, on-time, and late trials for linguistically congruent and incongruent conditions at the 450-500ms time window for left anterior channels (See Figure 8b). This analysis indicates that on average, mean amplitude was greater for out of key \((M = 1.54, SD = 3.97)\), than in key \((M = -0.82, SD = 3.17)\) trials \((t(24) = 2.23, p = 0.035)\) for rhythmically late trials. Moreover, there was a marginally significant three-way interaction among the factors of music, language, and rhythm at the 450-500ms time window \((F(2,46) = 2.91, MSE = 1790.20, p = 0.065)\) at the left posterior ROI. These results suggest that the processing of musical and linguistic syntax varies by rhythmic expectancy.

### 2.3 Discussion

Experiment 1 investigated the effects of rhythmic expectancy on the processing of musical linguistic syntax. This was conducted in an experimental paradigm in which the violations in the three factors were counterbalanced to occur unexpectedly and either independently of each other or simultaneously. Such presentation of stimuli allowed for us to test the theoretical frameworks of the SSIRH and DAT in our experimental paradigm, and how these frameworks may intersect in the interactive effects of co-occurring expectancy violations (i.e. Patel, 2003; Large & Jones, 1999).

Results from average amplitudes across time through comparison of all pre-critical, critical, and post-critical region ERPs showed significant reductions of the P2 (180-210ms after event onset) and N400 (450-500ms after event onset) components over time (Figure 4). These reductions indicate attenuation or habituation to the
auditory and visual stimuli late in each trial after critical region segments. Such attenuation, moreover, suggests that the cognitive and perceptual resources allocated for stimuli processing are reduced over time. Specifically, the reduction of the centrally-distributed N400 (see Figure 4) across time suggest a decrease in the allocation of attentional resources associated with the semantic processing and integration of the linguistic stimuli (Kutas & Hillyard, 1988), as well as in attentional reorientation (reorienting negativity) following the onset of each event (Schröger & Wolff, 1998).

As shown in previous research supporting the SSIRH, the prominent ERP components identified with violations in musical and linguistic syntax expectancy are Early Anterior Negativity (EAN), and a late positivity (P3a/LPC) (Patel et al., 1998; Koelsch et al., 2000; Koelsch et al., 2005; Loui et al., 2005; Steinbeis & Kolesch, 2008). The EAN was evoked in response to musically out of key chords and linguistically incongruent critical region segments (see Figure 5; Figure 6), corresponding in latency and prefrontal distribution maximally at the 180-210ms after event onset. Though it has been debated whether or not the EAN is lateralized for music and language (see Koelsch et al., 2005; Loui et al., 2005), with out of key chords eliciting a right-lateralized Early Right Anterior Negativity (ERAN) and linguistically incongruent sentences eliciting a left-lateralized Early Left Anterior Negativity (ELAN), the results from the current study do not show any patterns of hemispheric lateralization, but rather a strongly bilateral EAN. This early anterior negativity was not statistically significant for both musical and linguistic syntax expectancy violations at the tested ROIs, but was followed by several ERP
components that give insight into the temporal features of the cognitive processing mechanisms associated with syntactic anomaly detection. The lack of hemispheric asymmetry and statistical significance for the EAN maybe due to reduction of the component itself as a result of neurological interference in the simultaneous processing of musical and linguistic expectancy violations, thus this confounding variable will be further investigated in Experiment 2 (Koelsch et al., 2005).

In addition to the EAN, musical syntax expectancy violations elicited a P3a and a LPC. The P3a was significant at both midline anterior and midline posterior channels maximally at the 450-500ms time window, corresponding in latency and distribution to past research (Hillyard et al., 1973; Besson & Faïta, 1995; Luck, 2014). This positivity, moreover, suggests attentional reorientation to stimuli after detecting the deviation in musical syntax expectancy and that such reorientation to task-irrelevant musical out-of-key chords tax attentional resources (Munka & Berti, 2006). The LPC was significant across all posterior ROIs maximally at the 750-800ms time window, corresponding in latency and distribution to past research (Patel et al., 1998; Koelsch et al., 2005; Hahne & Friederici, 1999). This late positivity suggests memory-related context updating following the detection of the syntactic anomaly, reflecting decision-making based on implicit musical knowledge (Besson & Faïta, 1995; Loui et al., 2005). A similar, but not significant, LPC was seen following the EAN for linguistic syntax violations at the 750-800ms time window, suggesting memory-related context updating after detecting errors in syntactic garden-path sentences (Slevc et al., 2009). Together, the ERP components elicited by musical and
linguistic syntax violations replicate results and trends seen in studies supporting the shared resources hypothesis between music and language postulated by the SSIRH.

In addition to investigating the shared resources in musical and linguistic syntax processing, the current study aims to examine the effects of rhythmic expectancy on music and language processing and the extent to which rhythm is processed in a hierarchical manner similar to syntax processing. Though not significant, rhythmic expectancy violations yielded an anterior P2 enhancement at the 180-210ms, as well as a P3a enhancement (see Figure 7). The P2 enhancement suggests that violations in rhythmic expectancy elicit similar effects on this auditory ERP component by deviant stimuli in auditory sequences, suggesting that, at early time windows, rhythmic expectancy violations have a perceptual effect on high-level music and language processing (see Choi et al., 2014). Moreover, the P3a enhancement suggests that rhythmic expectancy violations elicit similar attentional re-orientation to stimuli as musical syntax expectancy violations, suggesting that rhythm may rely on similar processing mechanisms and neurological resources as music and language syntax processing (Fitch, 2013).

Follow-up interactive analyses further support the hypothesis that music and language processing rely on the rhythmic structure and presentation of stimuli. Specifically, the two-way interactions between music and rhythm at the left posterior ROI and between language and rhythm at the left anterior ROI at the 450-500ms time window reveal specifically how rhythm effects music and language processing. The enhancement of the P3a elicited by musical syntax expectancy violations by both early and late rhythmic violation conditions suggests that these temporal perturbations
tax the attentional resources associated with musical syntax processing, and replicate the attention curves associated with auditory deviant tone detecting tasks supporting the dynamic attending theory (Jones et al., 2002).

Despite these findings that suggest a hierarchical processing of rhythm in this experimental paradigm, as well as attentional effects of rhythm on music and language processing, these results are inconclusive in specifically pinpointing how rhythm relies on and affects the resources associated with syntax processing. Such limitations arise from the experimental design, in which violations in the three factors are less frequent, adding noise to the ERP data comparisons between violated and un-violated trials, and may occur simultaneously and do not allow for controlling for the confounding factor of neurological interference associated with such co-occurrence of different expectancy violations. Thus, Experiment 2 was conducted to investigate the independent effects of each factor, with violated trials occurring more frequently. This will allow for a better comparison of the ERP components elicited by rhythmic expectancy and musical and linguistic syntax violations, allowing for better insight on the extent to which rhythm affects and relies on these syntactic processing mechanisms and networks.

**3. Experiment 2**

Given the complexity in experimental design, it is possible that the three-way and two-way interactions and modulations to ERP waveform and amplitudes due to musical syntax, linguistic syntax, and rhythmic violations are due to the processing of multiple violations and such processing’s tolls on attending and development of
expectancy to task-irrelevant stimuli. Furthermore, the different experimental cells in each condition of Experiment 1 had different numbers of trials, in keeping with our own previous study (Jung et al., 2015) and with prior behavioral work that motivated these studies (Slevc et al., 2009). This results in more noise in the incongruent conditions for each factor. To eliminate these sources of possible confounds from Experiment 1, Experiment 2 independently tests the effects of each of the three factors on the ERP waveforms, while equating the numbers of trials in each expected or unexpected condition. Significant effects of musical and linguistic syntax, as well as rhythmic expectancy, on the ERP components of interest could validate the findings of the first experiment, and similar modulations on ERP components by each of the expectancy manipulations suggest that the domains of music, language, and rhythm tap into similar neural resources.

3.1 Materials & Methods

In Experiment 2, participants again read sentences broken down into segments paired with musical chords from a harmonic chord progression. Linguistic syntax expectancy was manipulated using syntactic garden-path sentences, musical expectancy was manipulated using chords that were either in key or out of key, and rhythmic expectancy was manipulated by presenting critical region segments early, on time, or late.

3.1.1 Participants

A new group of 10 undergraduate students (mean age (18.7 years), age range (18 – 21 years), SD = 1.06) from Wesleyan University participated in Experiment 2 for course credit. Recording errors resulted in the loss of data for 3 subjects, leaving
7 subjects for the final analysis. Of the remaining participants, all reported normal hearing, 3 participants (42.86%) reported having prior musical training, averaging 2.33 years (SD = 1.53), two (28.57%) participants identified as female, and five as male. Six reported that their first language was English, one was a native speakers of English and another language, and one had a language other than English as their first language. Other than English, participants’ first languages included Bangla. Two participants (28.57%) spoke more than one language. All participants had normal or corrected-to-normal vision and reported no history of psychiatric or neurological disorders. Informed consent was obtained from all subjects as approved by the Ethics Board of Psychology at Wesleyan University.

3.1.2 Materials & Stimuli

The same musical stimuli borrowed from Slevc et al. (2009) and parameters of stimuli presentation were used from the first experiment. However, in order to investigate the individual effects of each expectancy violations on ERP waveforms, four new experimental modules were designed. Each module contains 144 trials, with 12 sentences from Experiment 1 and 36 newly written sentences. Each trial consisted either of a musical syntax violation, a linguistic syntax violation, or a rhythmic violation. Each module contained: 48 trials with only linguistic syntax violations, 48 trials with only musical syntax violations, and 48 trials with only rhythmic violations (24 early and 24 late). Half of the sentences in a module were assigned a “yes” question, and the other half a “no.” The four different modules were synthesized to counterbalance which trials were assigned a “yes” or a “no,” and which of the rhythmically unexpected trials were assigned “early” or “late” violations.
3.1.3 Procedure

Similar to Experiment 1, participants gave informed consent and completed a short background survey before the experiment. Participants were again instructed to pay close attention to the sentences rather than the chord progressions, and the experimenter picked one of the 4 possible experimental modules at random. Segments appeared at a fixed IOI, and answered a yes-or-no comprehension question at the end of every sentence. Answering the comprehension question cued the next trial. The experiment lasted about 60 minutes.

3.1.4 EEG Recording

The same EEG recording parameters from Experiment 1 was used in Experiment 2. EEG activity was recorded from 64-channel silver chloride electrodes referenced to the ground channel (FPz) during recording. Electrode placement matched the International 10-20 system locations, and impedances were kept less than or equal to 5kΩ across all channels. All channels were continuously recorded and digitized with a sampling rate of 1000 Hz. Recordings took place in an electrically shielded, sound-attenuated chamber, and the experiment was conducted with the lights off.

3.1.5 Data Analysis

Behavioral comprehension question response were saved as text files from MaxMSP, and imported into Microsoft Excel and SPSS for statistical analysis. ERPs were averaged separately for expected and unexpected stimuli over the same time windows and same filters applied as the EEG data in Experiment 1. The same artifact rejection and segmentation and baseline correction criteria were applied as well.
Because the current experiment is still ongoing and there is still a limited pool of data, the reported data is based on preliminary results, and no statistical tests on the ERP components will be reported. However, plots of average ERP traces and trends in the data will be reported and investigated.

3.2 Results

3.2.1 Behavioral Results

On comprehension questions, participants performed significantly above chance in all conditions (overall $M = 85.49\%, s = 4.82$, two-tailed t-test against chance level of 50% correct: $t(6) = 19.46, p < .0001$). Because none of the participants performed below 50% on the comprehension questions, all participants successfully attended closely to the linguistic stimuli, while ignoring the auditory stimuli.

3.2.2 ERP Results

Visual analysis of grand averaged waveforms confirmed auditory N1-P2 complexes across all conditions in expected and unexpected conditions, as well as several ERP components elicited by unexpected, as opposed to expected conditions. The ERP components investigated in the current analyses are consistent in terms of latency with those seen in Experiment 1. These include an early anterior negativity (EAN) maximal at 180-210ms after an out-of-key chord or linguistically incongruent segment, a posterior positivity (P3a) maximal 450-500ms after out-of-key chords, a later negativity (N400) maximal 450-500ms and attenuates after pre-critical, critical,
and post-critical segments, as well as a posterior positivity (LPC) maximal 750-800ms after a linguistically incongruent segment.

3.2.3 Time Region Analyses

Preliminary analyses across the pre-critical, critical, and post-critical time regions replicate effects of habituation seen in Experiment 1. Visual inspection of the ERP waveform on midline anterior channel FCz, as well as the topographical distributions of ERPs at time windows of interest, confirm reductions of the P2 and N400 components across time (Figure 9). These reductions confirm attenuation or habituation at early and late time windows as musical and linguistic segments occur, as a result of reduced perceptual and cognitive resource allocation after the critical region of each sentence.

![Figure 9. ERP Component analyses across pre-critical, critical, and post-critical regions on FCz.](image)

2.2.4 Main Effects of Musical, Linguistic, and Rhythmic Violations

Visual inspection of the musically in-key vs. out-of-key waveforms at midline anterior channel FCz, as well as the topographical distributions of ERPs at time windows of interest, confirm the EAN and P3a seen in Experiment 1 (Figure 10). Moreover, the EAN appears to be more pronounced in Experiment 2 than in
Experiment 1. This confirms musical syntactic error detection in this experimental design, and also suggests attentional re-orientation.

Figure 10. ERP component analyses for musically in-key vs. out-of-key trials on FCz.

Visual inspection of the linguistically congruent vs. incongruent waveforms at FCz, as well as the topographical distributions of ERPs at time windows of interest, confirms an EAN and LPC elicited by linguistic syntax violations (Figure 11). Although the ERP plot on FCz does not seem to indicate this LPC, the topographical distribution of the ERP reveals the LPC and its posterior distribution. This confirms linguistic syntax error detection, and suggests memory-related context updating after linguistic syntax violations.

Figure 11. ERP component analyses for linguistically congruent vs. incongruent trials on FCz.
Visual inspection of the rhythmically early vs. on-time vs. late waveforms at midline anterior channel FCz, as well as the topographical distributions of ERPs at time windows of interest, confirm the enhanced P2 and P3a for rhythmically violated sentence segments (Figure 12). In particular, the P2 enhancement seems to be much more pronounced for late rhythmic expectancy violations. This further suggests shared similarities between rhythmic violations and musical syntax violations in eliciting a P3a.

3.3 Discussion

Experiment 2 was conducted as a control experiment to investigate the individual effects of expectancy violations in each factor on the event-related potentials associated with music and language processing. This was tested for in an experimental paradigm in which violations in each of the three factors were counterbalanced and occur independently of each other. Such presentation allows for the investigation the effect of rhythm on music and language processing as either its own syntactic structure that rely on similar resources, or taxes the attentional resources associated with auditory and visual stimuli processing (i.e. Fitch, 2013;
Large & Jones, 1999). Moreover, this control experiment allows for the investigation of possible confounds of simultaneous violation processing that led to patterns of neurological interference seen in Experiment 1 (Koelsch et al., 2005).

Results from average ERP amplitudes across pre-critical, critical, and post-critical time regions suggest a reduction of the P2 (180-210ms after event onset) and N400 (450-500ms after event onset) components over time, confirming the habituation to stimuli over time in the paradigms in both experiments (Figure 9). Moreover, the attenuation of these ERP components further suggest the reduction of the cognitive and perceptual resources associated with semantic processing and integration of linguistic stimuli and attentional reorientation following each segment in a trial, as seen in Experiment 1 (Kutas & Hillyard, 1988; Schröger & Wolff, 1998). Together, these ERP reductions over time, as well as the comprehension accuracy results, confirm the validity of this control experiment, and allow us to compare the ERP results of both experiments, and further investigate the theoretical implications of such results in relation to the SSIRH and DAT (Patel et al., 2003; Large & Jones, 1999).

Visual inspection of the preliminary ERP results for each factor in Experiment 2 replicate some of the key findings and implications of Experiment 1, and suggest further comparisons that give insight into the effects of rhythm on syntactic processing mechanisms. For both musical and linguistic syntax violation, the early anterior negativity (EAN) is much more prominent in Experiment 2 than in Experiment 1 (see Figures 10 & 11). This may suggest that the low amplitude of the EANs in Experiment 1 are a result of neurological interference in syntactic processing
resources associated with simultaneous processing of musical and linguistic syntax violations (Koelsch et al., 2005; Sleve et al., 2009). Moreover, preliminary results replicate the P3a seen for musical syntax violations and the LPC for linguistic syntax violations. This suggests that upon syntactic error detection, music and language syntax processing tend to diverge in later cognitive processes, with an attentional re-orientation for music, and a memory-related context updating effect for language (Munka & Berti, 2006; Besson & Faïta, 1995). These differences, however, may be a result of the experimental design, in which musical violations were task-irrelevant, whereas linguistic syntax violations were presented in a garden-path design (Slevc et al., 2009).

For violations in rhythmic expectancy, the preliminary results suggest a replication of the P2 and P3a enhancements seen in Experiment 1. Moreover, the P2 enhancement appears to be much more pronounced for rhythmically late than rhythmically early violations (see Figure 12). This further suggests that rhythmic expectancy violations have a perceptual deviant stimuli effect on high-level music and language processing (Choi et al., 2014), and also tax the resources associated with attentional re-orientation in musical syntax violation processing (Fitch, 2013; Munka & Berti, 2006).

Overall, the preliminary results help validate the paradigm and implications of results found in Experiment 1. By simplifying the experiment to investigate the independent effects of each factor, similar ERP components were elicited and, in some cases, more prominent, by expectancy violations in this experiment. However, the results of the current experiment are only preliminary, and based off of visual
inspection from a limited pool of data. Thus, more data must be collected to confirm the implications of these preliminary results.

4. General Discussion

The goal of the current study is to examine the effects of rhythmic expectancy on the processing of musical and linguistic syntax. Experiment 1 shows the ERP components elicited by violations in musical syntax, linguistic syntax, and rhythmic expectancy violations, as well as the interactive effects of rhythm on the ERPs associated with syntax processing. These data patterns confirm that rhythm affects the shared cognitive resources between music and language, and generally consistent with both the SSIRH (Patel, 2003) and DAT (Large & Jones, 1999). However, some of the follow-up analyses are inconclusive as to the exact nature of these interactions and the effects of rhythm on syntax processing. Although the interaction of music and rhythm at the left posterior ROI at the 450-500ms time window demonstrates that rhythmic violations enhance P3a amplitude associated with musical syntax violations, similar to attention curves seen in behavioral studies supporting the DAT (Jones et al., 2002), such patterns of enhancement were not seen between rhythmic and linguistic syntax violations (See Figure 8 & 9).

Moreover, within these interactions, it seems only rhythmically early violations significantly enhanced P3a amplitude for musical syntax violations, and rhythmically late violations enhanced ERP amplitude for linguistic syntax violations at the same time window. The reason for these differences in interactions is unclear, but it may arise from linguistic syntax violations being processed on a longer
timescale than their musical syntax counterparts in this experimental design. Alternatively, it may be a result of the complexity of our experimental design. Although these significant interactions and ERPs elicited by violations in each of the factors suggest the manipulations in this experimental design were effective, the inconclusive results may result from low power due to relatively low frequency of violations, as well as patterns of interference as a result of processing of simultaneous violations in the experimental design of Experiment 1 (Koelsch et al., 2005; Slevec et al., 2009).

Because it is possible results were consequent of complexity of our design, Experiment 2 simplifies the design by examining the independent effects of each factor of interest. Preliminary results of Experiment 2 show more prominent EANs for both linguistic and musical syntax violations, as well as trends suggesting replications of later ERP components, such as the P3a for music and LPC for language. This suggests that, despite the complexity of design in Experiment 1, the implications of these findings are valid. Specifically, it can be suggested that, in this experimental paradigm, that the neural mechanisms underlying musical and linguistic syntax structures appear similar in early processing (<300ms) but diverge later in time (>300ms). Thus, musical syntax violations tend to have more of an attentional re-orientation effect, as seen in the P3a (Besson & Faïta, 1995; Munka & Berti, 2006), and linguistic syntax violations tend to have a more memory-related effect, as see in the LPC (Slevec et al., 2009; Loui et al., 2005; Koelsch et al., 2005).

In both experiments, rhythmic violations elicited ERP components that share some similarities with auditory (P2) and musical syntax expectancy violations (P3a).
This may suggest that rhythm may manifest itself more readily in the auditory than the visual modality, and thus share more similarities with musical syntax than linguistic syntax in terms of resource allocation and processing mechanisms in this experimental paradigm (Fitch, 2013). However, it has been shown that rhythmic expectation has modality-general effects (Grahn, 2012; De Freitas et al., 2014; Escoffier et al., 2015; Escoffier & Tillman, 2008) and that rhythmic expectancy modulates ERPs associated with linguistic syntax processing (Schmidt-Kassow & Kotz, 2008). Therefore, the interaction between language and rhythm in this experimental paradigm must be further investigated, and follow-up research investigating the intersections of language and rhythm, independent of musical stimuli, must be conducted to investigate the central and modality-general effects of rhythm on cognition.

The investigation of the effects of rhythm on music and language processing is very relevant and important in the field of music cognition and cognitive science in general. Neuropsychological studies have shown that the link between linguistic and musical grammar processing can vary by musical and linguistic expertise (Gordon et al., 2015; Herdener et al., 2014; Fitrozy & Sanders, 2013; Jentschke & Kolesch, 2009; Sammler et al., 2009). Moreover, it has also been demonstrated that musical priming enhances music and language processing (Kentner, 2012; Przybylski et al., 2013). Thus, follow-up analyses should investigate the role of musical training as well as native English experience on the effects of expectancy violations on the ERP components of interest.
Overall, the current study demonstrates that rhythmic expectancy plays an important role in the shared processing of musical and linguistic structure. The investigation of shared processing of musical and language structure has been central to music cognition, as is relevant to investigating how rhythm affects attentional entrainment. The current study provides support for overlap in the processing resources for musical and linguistic syntax, and also suggests rhythmic perturbations tax attentional and syntactic resources. By offering a window into how perturbations of rhythmic and temporal expectancy affect musical and linguistic processing, the current results may translate towards a better understanding and possibly designing interventions for populations with speech and language difficulties, such as children with atypical language development (Jenetschke et al., 2008; Przybylski et al, 2013; Gordon et al., 2015a). We are surrounded by rhythms in our everyday auditory and visual environments, and understanding the neural underpinnings of rhythm and its shared resources with music and language can help us use these rhythms as a tool of cognitive development and repair.
Acknowledgments

I would like to thank my family for making my studies at Wesleyan University possible – my older brother, and best friend, Girim Jung, for always looking out for me, my father, Youngmin Jung, for the constant love and support, even through our unique family situation, and for the most beautiful woman in the world, my mother, Oksoon Jo, for her strength and loving sacrifices as a single immigrant mother to allow me to pursue my passions and dreams.

Professor Psyche Loui, and Professor Andrea Patalano for their support and guidance for my research in music cognition throughout my time at Wesleyan University, and for considering my research for my senior thesis.

Professor Roger Matthew Grant, Professor Ron Kuivila, and my anonymous third reader in the music department for also supporting my research and considering my thesis.

Professor Nadya Potemkina, Professor Ron Ebrecht, Professor Neely Bruce, Professor Paula Matthusen, and Professor Julie Ribchinsky for their inspiration and support of my studies of music and continued pursuit of classical double bass performance throughout my undergraduate career.

Ronnie Hendrix, Professor Suzanne O’Connell, and my McNair Scholars cohort for all the great memories and for being my second family at Wesleyan.

My super rad bandmates in Dark Circles, Chris Sailor, Delilah Seligman, and Maya Stevens, for sharing one of my most memorable experiences during my time at Wesleyan, and for helping me come to terms with my identity – you are literally my favorite people.
All the great people from the MIND Lab, past and present, who have supported me and my research throughout these years including: Janani Iyer, Emily Przysinda, Rachel Guetta, Wy Ming Lin, Kellyn Maves, Cameron Arkin, Keith Spencer, Maxime Bouvagnet, Tim Zeng, Min Cheol Lee, Sydney Lolli, Matan Koplin-Green, Mike Massone, Julian Basurto, Ari Lewenstein, and Aaron Plave.

Lastly, I would like to thank the friends and acquaintances I have made through my musical endeavors, as well as in the different student support groups, and political activism movements I have been involved throughout my time at Wesleyan.

Supported by startup funds from Wesleyan University, a grant from the Grammy Foundation to P.L., and the Ronald E. McNair Scholars Program to H.J. We thank all our participants.
References


