4-2017

The Convergence of Schenkerian Music Theory and Generative Linguistics: An Analysis and Composition

Michael A. Ciaramella
College of the Holy Cross, maciar17@g.holycross.edu

The Convergence of Schenkerian Music Theory and Generative Linguistics: An Analysis and Composition

Michael Anthony Ciaramella

College Honors Senior Thesis

Spring, 2017

Advisor: Professor Shirish Korde

Reader: Professor Chris Arrell

Michael Anthony Ciaramella
The Convergence of Schenkerian Music Theory and Generative Linguistics

Table of Contents

Michael Anthony Ciaramella
The Convergence of Schenkerian Music Theory and Generative Linguistics

Abstract

This thesis engages a purported connection between Schenkerian music theory and the Minimalist Program within generative linguistics both scientifically and creatively. The first chapter expounds the link between Schenkerian theory and the Minimalist Program which has been recently substantiated in a doctoral dissertation by Somangshu Mukherji at Princeton University and details the methodological framework for investigating musical structures within this paradigm. Chapter two presents three case studies including the opening phrase of Mozart's K. 332 Mvt. 1 piano sonata, and the tunes "Georgia on My Mind" and "Blue Bossa" in order to exemplify the aforementioned methodology and provide scientific evidence affirming this generative framework. Chapter three concludes with a creative investigation of the theoretical ideas which this thesis engages and consists of a string quart that draws upon the notions of music and language, and music as derived from a computational system.

Forward

The questions of the processes and forms that underlie musical structure which this thesis engages are a manifestation of a deeper question: What is the basis of human creativity? Scholars of all times, geographies and disciplines have debated this question and yet, it persists unanswered. Further, though the phenomenon of creativity seems almost impenetrable to inquiry, scholars have never tired of addressing it in novel ways. This thesis enters into this rich scholarly space by asking the question: How might the human mind generate an infinitely many unique and expressive musical structures from a finite set of musical inputs? By no means do I purport that I will arrive at a definitive answer for this question. Rather, it is my hope that journeying through this intellectual space will force me to think critically about the basis of music within the human mind, what makes it so meaningful to us, and what it means to be musical. Indeed, this inquiry marks my entrance into a life's worth of questions.

I was first exposed to the ideas in this thesis while a high school student taking music theory classes at Princeton University. Somangshu Mukherji was a graduate student at Princeton during my time there and an assistant for both of the classes I was enrolled in. Over several lunch conversations, Dr. Mukherji discussed his doctoral dissertation entitled, "Generative Musical Grammar-A Minimalist Approach" with me and opened my eyes to rich questions of how music and language interact with the human mind and what makes them so uniquely special to humanity. These questions stuck with me as I began my time at the College of the Holy Cross and grew in my mind under the supportive mentorship of many great professors, perhaps most notably my advisor, Shirish Korde, and reader, Chris Arrell, for this

thesis. Professor Korde has instilled within me an interest in world music and cross-cultural musical compositions and has always lent a patient ear for my explorations in this regard. Additionally, it was in Professor Arrell's music theory course that I was first exposed to the ideas of Heinrich Schenker and developed an interest in analysis. All the while, the questions raised to me while still a high schooler surfaced and resurfaced throughout these formative experiences at Holy Cross and have motivated me to pursue a thesis integrating my loves for composition, world music, and analysis to engage the questions Dr. Mukherji had presented me with.

This thesis is therefore indebted to the work of Dr. Mukherji and engages specifically with the ideas put forth in his Princeton doctoral dissertation. The first chapter expounds Dr. Mukherji's approach to generative musical grammar by engaging both his theoretical and methodological framework. Chapter two then presents three original case studies which investigate generative basis for three compositions from a diverse set of idioms. The chosen pieces not only manifest from different cultures, but also cover a timespan of almost two centuries. The goal of these case studies is to provide support for situating Dr. Mukherji's generative approach as an accurate reflection of an aspect of the human mind and of human nature. Finally, in Chapter three, I write a string quart inspired by the ideas of music's generative structure that were encountered previously in the thesis. By concluding in this way, I look to counterbalance my scientific and theoretical investigation of music with an artistic and creative encounter of the same ideas.

Chapter 1: Linking Music and Language

The first substantial attempt to link Schenkerian music theory to the Minimalist Program in generative linguistics was done by Somangshu Mukherji in his Princeton University dissertation entitled "Generative Musical Grammar- A Minimalist Approach." This thesis will draw upon Mukherji's work for its theoretical framework and methodology, which will be detailed in this chapter. To illuminate this theoretical framework and situate it within contemporary music theoretic discourse, Section A will develop two key distinctions that separate the application of the Minimalist Program from other research methodologies within the field of music theory. These distinctions are: formalist vs. functionalist, and generation vs. perception. The former in each are specific claims by Mukherji toward the application of the Minimalist Program to music. Exploring these claims will also problematize some of the other approaches to understanding musical structure and present application of the Minimalist Program as a potential avenue toward resolving some of these issues.

Sections B and C of this chapter provide the substantive basis for the connection between Schenkerian music theory and the Minimalist Program from generative linguistics, which was also put forth by Dr. Mukherji. The advances in generative linguistics through the Minimalist Program that make it suitable for connection to Schenkerian music theory are briefly reviewed, in order to then demonstrate the striking similarities between them and thereby substantiate the link between the two paradigms. The emerging program works within a formalist framework and seeks understand the psychological activity requisite to musical

activity and to characterize the basic mental processes by which humans generate musical

forms with infinite capacity for creativity and expression.[1]

Section A: Situating the Minimalist Program within Music Theory

Formalist vs. Functionalist

Music is an undeniable cultural artifact, since it is practiced in all known cultures

throughout the world. Therefore, any scientific investigation of music must acknowledge the

diverse kinds of musics found in cultures throughout the world and present an argument which

is able to account cohesively for their many differences and similarities. John Blacking, in his

book *How Musical is Man?,* arguably one of the most foundational texts of ethnomusicology

and certainly one of the most impactful, defined music as "humanly organized sound." Through

his emersion into the culture of the Venda, a tribal group in Western Africa, Blacking came to

observe that all people and cultures are capable of being musical, not just the select few who

have been fortunate enough to have had formal music training. Further, he recognized that

different musics manifest uniquely around the world and argued that cultural forces shape

these unique manifestations. On this basis, he advocates for the study of music by its *function*

in each society. He acknowledges that music seems to be a unique characteristic of human

behavior, but claims that to truly understand music, one must look at the culture from which it

was generated.[2]

---

[1] Mukherji, Somangshu. "Generative Musical Grammar---a Minimalist Approach." (PhD Diss. Princeton University 2014).
[2] Blacking, John. How Musical Is Man? Seattle: University of Washington Press, 1973. Ch. 1.

Blacking is arguing for a *functionalist* view of music.[3] This perspective has become ubiquitous among ethnomusicological literature and poses an obstacle to the study of music as something universal and uniquely human. If music is a uniquely cultural artifact, then it should not be subject to any natural law or scientific principle. For a functionalist, any attempt to map the origins of music, answer the question of why humans are musical, or explain what the capacity for music is, will inevitably end in the study of culture. This is an obstacle for any scientific study of music, because culture is known not to adhere to any natural laws, such as those found in evolutionary theory. This also poses a more direct obstacle for the connection between music and language. Music and language manifest in very distinct ways from a cultural perspective and serve very different functions in a sociocultural context. Thus, an entirely functionalist investigation of music will render it a completely separate phenomenon from language.

Dr. Mukherji, in his dissertation entitled "Generative Musical Grammar-A Minimalist Approach" addresses the issue of defining music in functional terms as Blacking does, rather than formal terms. Mukherji begins this discussion by acknowledging that while one of Blacking's goals in *How Musical is Man?* was to advocate for a functionalist perspective of music and claim that music is unique in each society, he was also speaking out against many ethnocentric and potentially racist conceptions of music commonly held in the Western tradition. Blacking recognizes that where in Western European culture is there a separation between performer and audience, in many other places around the world there is no such

---

[3] Proceeding discussion from Mukherji, Somangshu "Generative Musical Grammar-A Minimalist Approach" p. 12-22

distinction. In other cultures, such as the Venda studied in *How Musical is Man?*, music is an

inherently participatory practice among all members of a society, rather than just a

performance practice.[4] In recognizing this contrast, Blacking also calls into question the

distinction between "art" music, and "cultural" or "ethnic" music. The former, he argues, has

an association with high society and intellectual elitist culture. The latter, is associated with low

society and the lower class. Intrinsic to this distinction is also the value judgement that "art"

music is a superior practice to "ethnic" music by its association with high society. By extension,

this often leads to the ethnocentric perception of Western art music as superior to other

musical forms, on the basis of its complexity and association with high society. Blacking clearly

intended to topple this ethnocentric conception, in preference for equality and inclusion. This

is known in anthropology literature, and by extension in ethnomusicology, as an "emic" bias,

which overemphasizes certain aspects of a phenomenon to topple such prejudices.[5]

Blacking is not wrong in the assertion that all musics are equal and that value

judgements based on perceived superiority of certain societies is condemnable. On the

contrary, this is an admirable pursuit. However, this motive poses a problem for any claims

about the universal, and uniquely human faculty of music as a psychological phenomenon. By

overemphasizing the functional aspect of music, Mukherji claims that this approach becomes a

"study of differences."[6] This inherently poses a problem for a psychological theory of music,

which must look for what is shared throughout all humans, rather than the characteristics that

distinguish them. By painting each culture as entirely unique, and by extension, each kind of

---

[4] Blacking, *How Musical is Man?* Ch. 2
[5] Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 12-17
[6] Ibid. 17

music as unique as well, all we are left to do is discuss what differences there are between cultures and musics, without any standard for comparison. Mukherji therefore claims that such pursuit misses the opportunity to characterize perhaps the most powerful aspect of music: that it is a shared practice for *all* of humanity with some commonly held features.[7]

For Mukherji, one commonality of all musics is that they give humans the ability to express themselves and understand one another. This essential trait of music is also shared with language, as a uniquely human practice which allows for expression and understanding. The universal communicative aspect of music and language and their centrality to human expression led Mukherji to define these phenomena as fundamental aspects of *human nature*.[8] Therefore, music is irreducible to simply an artifact of culture, though he will argue culture is one factor that shapes it, and he argues for its study as part of human nature, as well as of human biology. Where Blacking acknowledges the unique human manifestation of music and goes on to explain its different realizations through culture, Mukherji focuses on what exactly the uniquely human characteristics of music are, hence focusing on its natural form, rather than its cultural function as Blacking does.[9] Additionally, he argues that understanding what music is in a human species-specific sense would benefit the study of music as a cultural artifact by creating a set of commonalities upon which the anthropological study of music might be based. Therefore, such study ceases to become simply a "study of differences," but rather a more fruitful cross-cultural investigation.[10]

---

[7] Ibid. 17-19
[8] Ibid. 11
[9] Ibid. 35
[10] Ibid. 17

Therefore, the application of the Minimalist Program to music is a *formalist* argument, rather than a *functionalist one*. Specifically, it treats music as an object of human biology, with a natural and universal structure underlying its seemingly unique manifestations across culture. It will claim that music, "is based on an inner model, that governs its external, individual characteristics," taking the internal form of a finite hierarchical and recursive computational system embedded within the human mind, manifesting externally as the infinitely unique array of sounds which we hear as music.[11] Such a claim makes music a human psychological reality, and an indispensable aspect of human nature. The Minimalist Program in generative linguistics likewise treats language as an aspect of human biology with a singular internal form, which manifests in the same ways that Mukherji, proposes music does.[12] On these grounds, he connects the well-defined research program known as the Minimalist Program to music to generate a psychological theory of music.

There is one final point relevant to the formalist vs. functionalist argument. This is the difference between musical (and linguistic) *competence* versus *performance*. In suggesting musical and linguistic behaviors to be aspects of human biology, with which all humans are naturally endowed at birth, the Minimalist Program is claiming that all humans are born with musical and linguistic *competence*. This means that all humans minds are encoded with the specific finite computational system which allows the hierarchical and recursive operations necessary to produce and comprehend music or language. This says nothing, however, about how music and language manifest externally in terms of sounds, word order, and other

---

[11] Ibid. 143
[12] Ibid. 138-144

characteristics, other than that they manifest from the internal computational system that

defines music and language. These external characteristics are defined as musical and linguistic

performance.[13] Thus, someone working within the Minimalist paradigm, as Mukherji does,

would respond to Blacking's observation that the music of each culture is unique by saying that

musical *performance* allows for its unique characteristics. The unique ways these sounds come

together, however, is governed by the hierarchical and recursive computational system that

manifests internally within the human mind.[14] It is this aspect of musical and linguistic

competence which the Minimalist Program claims is a universal human trait and attempts to

characterize. This distinction between performance and competence will also be relevant to

the other three distinctions which this chapter makes.

Generation vs. Perception

Given its focus on linguistic competence, rather than performance, the Minimalist

Program is specifically focused on the mental process by which linguistic structures are

generated with infinite creativity from the finite material of the mind's computational system.

As such, the Minimalist Program resists the notion that the perceived sounds that make up

language (or by extension, music) may be *reduced* to some abstract form, but rather begins

with the fundamental abstract units of language (or music) and describes the mechanism by

which such material is formed into a recognizable language, allowing for the infinite creativity

seen in language and by extension, music. To be described more in Section B, the mechanism

by which the Minimalist Program claims language is constructed is called "Merge," in which

---

[13] ibid. 33-39
[14] Personal communication with Dr. Mukherji.

lexical linguistic units are combined into larger units called phrases. These phrases may then be combined into larger units, forming language we hear, read, or see. Two constraints, however, are applied to the products of Merge before language is realized. These are that they conform to a *phonetic form*, which entails that the product is able to be articulated via our sensorimotor system; and *logical form*, which requires that the product passes some conceptual-intentional system within the brain so that it has a meaning that is intelligible. The Minimalist Program is specifically *generative* in orientation, as opposed to *reductive*, since it works from the bottom up, showing how our minds create language from the natural biological linguistic system which we are endowed with.[15]

An extension of the work done by the music theorist Henrich Schenker will provide a foundational link between the Minimalist Program for language and for music. The most important basis of this connection is the claim that Schenker's music theory, like Chomsky's linguistics, is *generative* in nature, rather than reductive. Before this claim is defended, it must be noted that any literature which uses Schenker in a scientific context, or seeks to apply his theory to anything outside of common practice Western tonal music must acknowledge a sort of "neo-Schenkerian" approach as opposed to a strict and traditional sense of Schenkerian theory. Schenker was an Austrian music theorist living from 1868 to 1935 who focused exclusively on the works of the "Germanic master composers," which limited his scope to the analysis of essentially the Germanic composers from Bach to Brahms chronologically. His analytical approach was aimed at highlighting the mastery of certain group of composers within

---

[15] Berwick, Robert C., and Noam Chomsky. Why Only Us: Language and Evolution. Cambridge, MA: MIT Press, 2016. Pg. 96-102

the Western tonal tradition. Additionally, Schenker claimed that these so called "master

composers" were the only ones capable of creating music of such skill. Inherent to his work,

therefore, was an elitist and nationalist bias which limited both the scope of his analyses and

theory, as well as his reception in the music community. Thus, Schenker was blinded by his

biases to the true power of his theory. This is a stance adopted by all of his students who came

to the United States in the post-war era and vehemently condemned his biases and politics, but

advocated for and taught his technical theories in the country. Nevertheless, it may be argued

that the process that Schenker set out to depict through his analyses can be viewed as the basis

of a generative theory of music. In essence, he desired to characterize the process by which

master composers intuitively built upon a single form and generated unique and genius musical

creations with seemingly infinite creativity.[16] If we strip away Schenker's reservation that only

a select few were capable of such a process, what remains is the claim that musical structures

are generated with infinite creativity upon a single frame. This idea is also shared with

generative linguistics and provides the most fundamental connection between the two

disciplines.

Even removing the nationalist and elitist restrictions from Schenker's theories and

allowing them to be applied to other cultures and time periods, it is a further extension to make

any psychological claims about music from Schenkerian theory. Schenker's theories were

meant to showcase natural the genius of the few master composers, as well as provide musical

performers with an analytical tool to aid their performance. By performing a Schenkerian

---

[16] The above discussion from many personal communications with Professor Shirish Korde and Dr. Mukherji as well as Professor Chris Arrell's Music Theory 3 course.

analysis of a piece of music, one of Schenker's intentions was for a performer to be clued into

the genius of the composer's work, in the hopes that this knowledge would allow for a more

virtuous performance of the piece.[17] This interpretation of his theory is known to some as the

"conservatory Schenker," which functions as an analytical poetic device for performance.[18]

Regardless, it is apparent that Schenker did not mean for his work to hold any psychological

implications. The psychological extension of Schenker's work emerged much later in the 20th

century, almost concurrently with advancement in the field of Generative Linguistics, because

of the many similarities between the two. The most famous attempt is Lerdahl and

Jackendoff's famous *A Generative Theory of Tonal Music*, which seemed to be the peak of

interest in the connection between Schenkerian music theory and generative linguistics.

Mukherji engages this text in detail in his dissertation, since it purports to address the same

issue of generative musical grammar. He ultimately concludes that their attempt failed. This

will be discussed at greater length to follow.

Returning to Schenker, it is specifically the generative conception of his analytical

paradigm that has drawn interest for a connection to linguistics. In Schenkerian theory, one

analyzes a piece of music by breaking it down into three layers, the foreground, middleground,

and background, to reveal the fundamental abstract structure upon which the composition is

constructed. This most foundational layer within a Schenkerian analysis is the background,

which reveals what Schenker called the "*Ursatz*", or "fundamental structure." The *Ursatz* is the

fundamental harmonic structure upon which the entire composition is built, in the form of a I-

---

[17] This stance is taken by other advocates of analysis as well, such as Kofi Agawu. See Agawu,
Kofi "How We Got Out of Analysis, and How to Get Back In" *Music Analysis,* 2004: 267-286
[18] Mukherji, "Generative Musical Grammar-A Minimalist Approach" pg. 159

V-I progression. Accompanying the *Ursatz* at the background level is the foundational voice

leading structure of the piece, known as the *Urlinie*, or "fundamental line." The *Urlinie* can take

three separate forms at the background level, each a descending line, which include scale

degrees (3-2-1), (5-4-3-2-1), or (8-7-6-5-4-3-2-1). Upon the foundation of the *Ursatz* and

through the *Urlinie*, the master composer builds a composition by employing voice leading

principles originating in species counterpoint. Each harmonic area of the *Ursatz* is "composed

out," which Schenker termed *Auskomponierung*, which adds layers upon the background in a

hierarchical and recursive manner, eventually leading to the surface music heard by the

listener.[19]

Mukherji believes that Schenker specifically intended this process of "composing out"

the *Ursatz* to be *generative* in nature and uses this as the basis for the connection of

Schenkerian music theory to the Minimalist Program. Schenker claimed that the *Ursatz* arises

from the "*Der Naturklang*," or "chord of nature," which is the harmonic sequence that arises

above any fundamental pitch. The basic triadic *Ursatz* of Schenkerian theory is therefore based

more fundamentally on natural acoustic laws, to which the human mind is attuned.[20] Thus, the

distinctive ability of the master composers was their ability to connect with the natural form of

the *Ursatz*, and build a composition upon these roots. This will provide the basis for a

psychological extension of Schenker since for him, music is fundamentally dependent on the

ability of the human mind to connect with the most basic material for music (the harmonic

tones which form the *Ursatz*) and build infinitely many unique structures upon this finite

---

[19] Ibid. 125-128
[20] Ibid. 142

fundamental form. A psychological extension of Schenkerian theory within the paradigm of the

Minimalist Program, however, will modify Schenker's conception of the most basic musical

material slightly. Instead of claiming that the *Ursatz* is the most fundamental characteristic of

music, arising from the chord of nature (the harmonic series), this approach will take the *Ursatz*

as a cultural artifact within the Western Classical idiom that arises from a more basic set of

universal musical material. To be discussed more in Sections B and C, this approach will claim

that a finite set of "chord-like" atoms within a universal musical lexicon is the most

fundamental musical material and innate within the human mind. A process called "Merge", in

which lexical items are combined in binary arrangements results in structures like the *Ursatz.*

The idea of the *Ursatz* also renders Schenker's theory distinctively generative. Rather than

beginning with the surface, a composer connects unconsciously with the most fundamental

abstract structure of music (for Schenker, the *Ursatz*, and for this approach, the combination of

the musical lexicon and Merge), realizing the surface in infinitely creative ways based on a finite

abstract structure through the principles of voice leading and counterpoint. It was the ability to

truly connect with the *Ursatz* as the fundamental material for composition, which Schenker

reserved for so few Germanic composers, that will become the psychological underpinnings of

the generative "scientific Schenkerian" theory.[21] Much more will be said about Schenker's

theory, its correlations with the Minimalist Program, and its overall psychological extension.

For now, it is sufficient to recognize that Schenker can be interpreted as *generative* in nature,

rather than reductive.

---

[21] ibid. 128-138

The first significant attempt to link music and language within the paradigm of generative linguistics came from Fred Lerdahl and Ray Jackendoff in 1983 in their work entitled "A Generative Theory of Tonal Music." With Lerdahl being a Schenkerian music theorist in addition to a composer, and Jackendoff being a student of Chomsky, the work sought to link the principles of Schenker to the advancements of generative linguistics at the time. These authors ultimately failed in their attempt at a generative theory of tonal music, one which they even acknowledged in their 1983 work by recognizing it as a theory of music perception.[22]

Mukherji's dissertation, completed in 2014, is a far more recent attempt to connect music to language within the generative linguistics paradigm. His attempt is more fruitful than that of Lerdahl and Jackendoff both because he works within a different, more generative conception of Schenker, and because the field of generative linguistics has advanced leaps and bounds beyond where it was in 1983 with the advent of the Minimalist Program. Thus, these two distinctions separate his dissertation drastically from Lerdahl and Jackendoff's in 1983. Mukherji addresses the shortcomings for Lerdahl and Jackendoff's *A Generative Theory of Tonal Music* in becoming a true application of generative linguistics to music.

The first shortcoming of *A Generative Theory of Tonal Music* is its inability to account for musical rhythm internally. In their work, Lerdahl and Jackendoff recognize rhythm as a completely separate entity to tonal harmony, which is governed by different rules unrelated to language in any way. As such, they develop a complicated system for the reduction of metric structures in an attempt to show a generative origin separate from tonal structure. Their

---

[22] Lerdahl, Fred, and Ray Jackendoff. A Generative Theory of Tonal Music. Cambridge, MA: MIT Press, 1983. p. 5

theory of rhythm employs the process of grouping, which is common to many areas of human

cognition. Grouping entails that the human mind "chunks" a series of elements perceived in

linear order into a larger "group." Such groups reduce the number of unique elements held in

the working memory (which is only capable of holding on average 7 units at any one time), and

allows for easier recall from short term memory into the working memory.[23] Lerdahl and

Jackendoff proposed a hierarchical arrangement of such groups that conforms to a set of

"preference rules" which govern the structures that are likely to appear and those that are not.

Based on the interaction between rhythmic grouping and their proposed preference rules, they

created a system of rhythmic reduction which occurs concurrently, but separately from tonal

cognition.[24]

The main reason Mukherji rejects Lerdahl and Jackendoff's theory of rhythm lies in

their "externalist" argument treating the tonal and rhythmic aspects of music separately.

Armed with the Minimalist Program, Mukherji rejects this externalist delegation of musical

rhythm and employs "phonetic form" level of representation found within the Minimalist

Program to account for rhythm internally along with other pitch-based aspects of music. To be

discussed more in Sections B and C, the binary character of Merge necessitates that rhythm is

created because it requires the ordering of linguistic (or musical) units in space, whether this

space is metered or not. For someone working within the Minimalist paradigm, rhythm must

therefore be grouped internally along with pitch, since a single mechanism leads to their

---

[23] See Alamia et al., 2016 for an argument that the language centers of the brain also operate by this "chunking" mechanism.
[24] Lerdahl and Jackendoff, *A Generative Theory of Tonal Music*, Chapters 1-4

construction. Thus, Mukherji argues for an "internalist" approach to rhythm.[25] The phonetic

level of representation just ensures that the ordering Merge created is capable of articulation

by the sensorimotor system in some way. It is therefore at this level that rhythm is realized for

the Merge output with the goal of satisfying the constraints of our sensorimotor system. Only

products which meet the demands of this level, as well as the "logical form" level via a

conceptual-intentional interface, will emerge as the music we hear.[26] More will be said about

this in Section B. The important point at this step is to recognize that through the

advancements of the Minimalist Program, there emerges an answer to one of the fundamental

difficulties that led to the failed generative account of music by Lerdahl and Jackendoff in 1983.

The second shortcoming which Mukherji acknowledges relates to Lerdahl and

Jackendoff's famous "preference rule system," on which their theory of both rhythm and pitch

is predicated. Lerdahl and Jackendoff acknowledge in their first chapter that a major

motivation of their work was to produce an empirically verifiable theory of tonal music, as this

was becoming a trend in music theory and specifically music theory that related to cognitive

science.[27] Their focus on empirically verifiable results inevitably forced them to look to musical

surface structure (i.e. the notes that are actually heard in music) as the basis of their theory.

Specifically, their focus was on how listeners parse through information at the musical surface

(which they explicitly heard) in order to arrive at a preferred structural description in their

mind. The listener-based and surface-based approach of Lerdahl and Jackendoff led them to

derive four categories of "preference rules" which listeners employ to arrive at a sense of

---

[25] Mukherji, "Generative Musical Grammar-A Minimalist Approach", p. 135-138, p. 604-611
[26] Berwick and Chomsky, *Why Only Us*, pg. 96-102
[27] Lerdahl and Jackendoff, *A Generative Theory of Tonal Music*, pg. 5

musical form. Their preference rules were based separately on grouping structure, metrical

structure, "time-span' structure (essentially a combination of the previous two), and

prolongation structure. Their theory necessitated that musical surface must simultaneously

meet the preference constraints proposed in each of these four areas simultaneously. Thus,

the simultaneous set of constraints governing musical surface limited the potential

interpretations listeners were allowed, allowing listeners to come to consensus at a preferred

structural interpretation of a piece.[28] Researchers could therefore employ these sets of

preference rules to empirically verify whether the constraints imposed by the theoretical

system conformed with listeners' actual interpretation of a piece.

Mukherji claims that the constraint-based approach to musical surface at the heart of

Lerdahl and Jackendoff's generative theory rendered it actually closer to a theory of music

perception, rather than *generation*, and also specifically anti-Chomskyian. A generative theory

would begin with the fundamental abstract form of music (which is the finite computational

system within the human mind based on the lexicon and Merge) and detail how the mind arises

at an infinite number of unique structures, rather than how a listener satisfies multiple

constraints to arrive at a consensus for structure. Additionally, this separates the processes

involved in composition from listening, since preference rules apply to listeners' interpretation,

not composers' creation. All of this renders the approach by Lerdahl and Jackendoff actually

anti-generative and by extension, anti Chomskyian, since Chomsky's generative linguistics

specifically views the creator and listener of linguistic structures as abiding by exactly the same

---

[28] Ibid. 1-8

mental process.[29] The Chomskyian generative paradigm begins with an abstract computational

form, rather than a constraint-based preference system. Mukherji therefore separates his

attempt to apply the Minimalist Program of generative linguistics from Lerdahl and Jackendoff's

attempt on the grounds that it truly is a generative theory which views the process of

composing and listening as being governed by the same basic mental faculty, and that this

faculty builds on a simple computational mechanism with a finite set of resources, in order to

create infinitely many unique outputs.[30]

Fred Lerdahl's later publication, *Tonal Pitch Space*, written in 2001, builds upon his work

with Ray Jackendoff in *A Generative Theory of Tonal Music*, and is closer in relation to a true

theory of generation than of perception.[31] In *Tonal Pitch Space*, Lerdahl constructs an

algorithm to govern the relative distance between tonal areas as perceived by humans. Closely

related tonal areas, like the subtonic in relation to the tonic, or the predominant in relation to

the tonic, are perceived as being "closer" than other tonal areas. The farthest distance within

an octave in this theory is the tritone, and is subsequently the least stable in relation to the

tonic. Lerdahl quantifies these relations and develops an algorithm to govern overall

relatedness given a set of tones. His theory can therefore account for the stability of related

keys quantitatively and predict listener perception of stability based on the quantitative result.

As such, he is capable of explaining why it is common in Classical music to modulate to the

---

[29] In fact, Berwick and Chomsky *Why Only Us: Language and Evolution* Ch. 1 actually suggests that language evolved as primarily an internal mechanism for thought. The externalization of language is hypothesized to have occurred after the mental computational system was already present upon evolutionary refinement of the vocal apparatus.

[30] Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 167-169

[31] The following paragraph focuses on the general argument put forth in Lerdahl, Fred. Tonal Pitch Space. New York: Oxford University Press, 2001.

dominant or the subtonic, since they represent a stable relationship for listeners. He is also able to account for tonal ambiguity in late Romantic music as well as music of the twentieth century. These genres explore "tonal spaces" which are farther away in relation to each other, and thus produce a different effect on the listener. Lerdahl's *Tonal Pitch Space* is closer to a generative theory because he seeks the intrinsic musical properties of chords that create the different pitch spaces. With this orientation, he is concerned more with how musical structures are generated and why they are constructed in the ways we observe, rather than how they are received perceptually.   Where *Tonal Pitch Space* seems to answer many questions regarding why listeners prefer certain sonorities over others by looking to the intrinsic properties of their construction, and how certain groups of pitches creates vastly different effects on listeners than others, it still does not answer the question of why music is generated in this manner to begin with. In essence, it characterizes Merge outputs and theorizes as to why these outputs are generated, but does not address the larger system generating these outputs. Thus, *Tonal Pitch Space* may be useful for classifying musical structures once they exist, but it fails to adequately consider why or how they come into existence at all. The extension of the Minimalist Program into music will provide answers to such fundamental questions.

Issues surrounding music generation versus perception also weave themselves into research in cognitive psychology and neuroscience. Concurrent with much of the theoretic research into the generation of musical structures and their perception, has been an effort to verify many of the claims found in these theories on neuroscientific grounds. One of the most significant claims for the Minimalist Program and its extension to music is that both are governed by the same hierarchical and recursive structure. Therefore, into the twenty-first

century, there has been an effort to examine the effect of syntactic changes to musical or

linguistic structure and their affects in known language-focused areas of the brain.[32] Such

research often measures the frequency of "event-related potentials" or specifically ERP 600's

(600 stands for the 600 ms time interval between the stimulus presented and the resulting

potential in the brain) which are known to register in the brain when an error in syntax is

perceived.[33] For example, Patel et. al. 1998 present listeners with a harmonic progression that

would commonly be found in common practice Western tonal music and examine brain activity

and ERP 600 responses.[34] Then, in the experimental group, researchers would alter one chord,

or change the order of the chords in some way that would result in a different underlying

syntactic structure. These alterations have been shown to cause ERP 600 peaks within the

regions of the brain traditionally considered to be involved with processing linguistic syntax.

Research of this kind has been replicated in a number of different settings and with different

genres and has generated similar results, leading researchers to infer that the brain is in some

way attuned to changes in musical syntax. This has led to a number of theories relating music

perception to syntax, some claiming a connection to linguistic syntax, and others claiming

---

[32] See Sammler et. al. 2011, Patel and Gibson 1998, Maess et. al. 2001, Kunert et. al. 2015, Koelsch et. al. 2007, Koelsch 2006 for some examples of this methodology

[33] Patel, Aniruddh D. *Music, Language, and the Brain*. New York: Oxford University Press, 2010. Pg. 273-276

[34] Patel, Aniruddh D., Edward Gibson, Jennifer Ratner, Mireille Besson, and Phillip J. Holcomb. "Processing syntactic relations in language and music: An event-related potential study." *Journal of cognitive neuroscience* 10, no. 6 (1998): 717-733.

apparent similarity between linguistic and musical syntax, but ultimately denying their

similarity.[35][36]

Overall, while these kinds of studies are helpful for determining the structures within

the brain involved in processing music, and for making observations about the nature of music

processing and representation in the brain, they are incapable of providing the level of

precision required to generate a theory of music encompassing its natural form, construction,

and implementation in the brain, as a generative theory connected to the Minimalist Program

would do. As a result, these studies address similar questions regarding listener perception

found in *A Generative Theory of Tonal Music* and *Tonal Pitch Space*, except from a

neuroscientific standpoint rather than a cognitive one. Therefore, they are open to similar

critiques from someone working within the Minimalist paradigm since they overlook the

fundamental computational structure of music, which is indispensable to its representation in

the mind and implementation in the brain's hardware. Their results have a certain degree of

ambiguity which necessarily lead to conjecture regarding the true nature of music. For

example, while it might be said that a registered ERP 600 peak in response to an altered musical

chord reflects a noted change in syntactic structure as represented from the deeper

computational system, it has also been proposed that this response is rather a more general

response to a perceived error in the music that may arise from any number of reasons, such as

a familiarity with the idiom generating a sense of expectancy. This line of thinking has led some

---

[35] Patel, Aniruddh D. *Music, Language, and the Brain* p. 239-300 covers this in great detail in a chapter entitled "Syntax"

[36] Dr. Mukherji also discusses the relation between neuroscientific and cognitive work to the Minimalist framework in "Generative Musical Grammar-A Minimalist Approach," p. 61-74

researchers to propose that the region of the brain in which these ERP 600 peaks occur is overall responsible for the interpretation of fine-tuned structures, and that more general errors register the ERP 600 peak.[37] Further, designation as an "error" resulting in an ERP 600 may also be seen to result from the repeated exposure of an individual to this kind of stimulus, which may have led to a sense of expectancy regarding the perceived "correct" form of the music at hand. Development of this sense of expectancy does not necessarily have to arise from a more abstract computational system, but rather may be developed linearly as a set of probabilities based on perceived pitches which may be received entirely from one's culture.[38,39] For example, consider how an avid blues listener may perceive a IV chord at the end of the twelve-bar form to be an "error" and register an ERP 600 peak. This may represent an alteration of the computational system, or the Merge output, or it could simply arise from a set of known probabilities relating to the blues, which were acquired linearly by avid listening.

The Minimalist Program seeks to avoid such ambiguity by addressing the question of why linguistic structures (or musical structures) are generated, rather than characterizing them once they exist. To do so, it adopts a specific philosophical paradigm of the human mind as the basis of its theory which is capable of fully representing music as a universal and natural aspect of humanity. This paradigm views the mind as an information-processing machine with three levels of operation necessary to describe any phenomenon. These levels are defined as

---

[37] Alaima et. al. 2016 is a recent example of this.

[38] Mukherji, "Generative Musical Grammar- A Minimalist Approach" pg. 171-177

[39] Berwick and Chomsky *Why Only Us: Language and Evolution* Ch. 2 entitled "Biolinguistics Evolving" also discusses an analogous strain of thinking found in linguistics that argues for a linear, probability based generative process for language. They argue against this stance in great detail.

computation, algorithm, and implementation.[40] Computation is the most fundamental and

abstract layer dealing solely with the most basic inputs and outputs of the mind. Minimalist

Program is specifically concerned with characterizing the mind's linguistic (or musical)

computational system by characterizing the lexicon (inputs) and Merge (the mechanism by

which outputs are realized). Products of the computational system are then applied at the

algorithm level, which is analogous to the software of a computer. At this level, the mind

receives input from the computational system and manipulates it in some way to produce an

output. The Minimalist Program claims that the "phonetic form" and "logical form" necessary

to representing language are aspects of this level. The final level is implementation, which

describes the actual neurological structures necessary to construct the system. Broadly, this

paradigm claims that one cannot simply look at any level individually, but rather how the

principles of each level connect with and affirm each other in order to fully understand a

phenomenon.[41] David Marr is one of the founding scholars of this paradigm and applied it to

human vision. The above description of this very intricate cognitive theory is extremely brief,

but interested readers can find more on this in Marr, 2010.[42]

The most important takeaway from this section on generation versus perception is that where

theories of music perception often describe structures once they exist and characterize their

---

[40] Marr, David. Vision: a computational investigation into the human representation and processing of visual information. Cambridge, MA: MIT Press, 2010

[41] Berwick and Chomsky, *Why Only Us*, Ch. 2 entitled "Biolinguistics Evolving" addresses this process in great detail and argues that its simplicity and efficiency also make its evolution feasible from both the perspectives of bioenergetics and selection.

[42] Marr, David. Vision: a computational investigation into the human representation and processing of visual information. Cambridge, MA: MIT Press, 2010.

form, the generative theory of the Minimalist Program seeks to answer the question of why

and how they exist at all. Taking Marr's paradigm of the mind, the goals of the work of Lerdahl

and Jackendoff, as well as much of the neuroscience and psychological research dedicated to

music detailed above, are actually to characterize the implementation or algorithm level of

music, without addressing the computational level. This will always fall short in truly coming to

a deep understanding of a phenomenon, since it is an essential pillar of this paradigm that the

levels interact and affirm each other at a theoretical level. Therefore, the Minimalist Program

seeks to work from the bottom up, beginning at the computational level, and moving up from

there to algorithm and implementation to truly characterize music, as opposed to beginning

with a loose sense of perception and characterizing what already exists.

Section B: An Overview of the Minimalist Program within Generative Linguistics

In this section, I will review the Minimalist Program from generative linguistics in

preparation of establishing its connection to Schenkerian music theory. In doing so, I will not

only characterize the mental computational system defined within the Minimalist paradigm,

but also put summarize the scientific methodology which one would use for an investigation of

human musicality within this paradigm. Additionally, I also draw substantially upon content

from this section as inspiration for the string quartet associated with this thesis entitled "An

Exploration of Musical Syntax."

The original motivation of the Minimalist Program was to streamline the complex

theoretical framework that already existed within generative linguistics into its most

fundamental and irreducible form.[43] This motivation arose partially out of concern that the field

of generative linguistics had gotten too complicated and theoretically convoluted to be

biologically feasible. It appeared as if linguists were over-generating grammar rules governing

the so-called "universals" of linguistic structure in ways that seemed to evade intuitive

sensibilities about psychological and scientific realities. The restrictions became so significant

that neurolinguists, psychologists, and evolutionary scientists began to notice that the cognitive

demand of parsing through the many linguistic syntactic rules subconsciously would likely

require the brain to exhaust all of its energetic resources solely on this task.[44] For this reason,

the ideas of linguistics came under fire.

Concurrently, linguistic theory became suspect from an evolutionary perspective.[45] The

selection advantage conferred by language acquisition is seemingly indisputable, but the

process by which language was selected for is a more challenging evolutionary question. It is

known that single mutations are easier to introduce to a population than multiple concomitant

mutations. Therefore, for language to be feasible from an evolutionary standpoint, it is either

the case that language arose quickly by a single, or perhaps very few, mutations that altered

the neural hardware, or that many necessary neural mutations made language adoption occur

very gradually over time. If language acquisition occurred gradually in humans, it would be

---

[43] See Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 243-317 for an overview of the history of generative linguistics leading up to the Minimalist Program. Page 281 is particularly important in this regard.

[44] Berwick and Chomsky *Why Only Us: Language and Evolution* Ch. 1 entitled "Why Now?" discusses this point in detail.

[45] The following discussion of the evolutionary pressures on language development within this paragraph is also found in Berwick and Chomsky *Why Only Us: Language and Evolution*, specifically Chapters 1-3.

expected that archeological and anthropological findings would reflect a gradual adoption over

the course of millennia of things like symbolic culture and other similar phenomena used as

proxy measures for linguistic ability. These findings, however, are not observed. Instead, it

appears as if language developed rapidly on scales of evolutionary time. This would appear to

fly in the face of Darwinian selection, in which mutations conferring an advantage are selected

for gradually, especially when involving multiple concurrent mutations. A plausible explanation

to reconcile this apparent contradiction, however, is that language was in fact the result of one

or very few mutations to the neural hardware. This was recently theorized by Chomsky and

Berwick in their 2016 book entitled *Why Only Us: Language and Evolution*. Perhaps, a single

alteration within the neural hardware endowed the human mind with the computational

system necessary for language. This would only be possible, however, if this computational

system was extremely efficient and simple, as to require few necessary mutations. This is

relevant to establishing a theoretical framework for linguistics because it supports the idea that

linguistic theory must have a simplicity and elegance that reflects its evolutionary adoption.

With the ideas of biological feasibility and scientific elegancy in mind, the Minimalist

Program claims that language can be attributed to a single computational operation, Merge,

which acts upon a finite set of "word-like" inputs with definable properties, which can generate

linguistic structures with infinite complexity and creativity.[46,47] This computational system is

---

[46] In what follows, see Berwick and Chomsky, *Why Only Us: Language and Evolution* p. 96-102 for a brief the linguistic computational system and the logic behind its efficient construction, as summarized within this paragraph.

[47] In what follows, see Mukherji "Generative Musical Grammar-A Minimalist Approach" p. 315-327 for a more detailed description of the linguistic computational system in Minimalist terms, as summarized within this paragraph.

said to exist for the sole purpose of satisfying the biological constraints to language in order to

generate interpretable structures with maximal efficiency. These constraints are that the

emerging structures meet the demands of a "Logical Form" which passes through a conceptual-

intentional interface within the mind and a "Phonetic Form" which mandates that it is capable

of articulation by the sensorimotor system in some way. Figure 1.1 below illustrates a mental

map of the linguistic computational system with each component labeled. Shown in the figure,

Merge is the central mechanism mediating the transition from linguistic lexical items to

intelligible linguistic structures. Thus, the ability to carry out Merge operations has been

proposed as the fundamental characteristic of human language that separates it from other

animal communication systems. Berwick and Chomsky also propose the acquisition of the

Merge ability to be the fundamental evolutionary change accompanying language's rapid

adoption historically.[48] Importantly however, each aspect of the computational system is

inseparable from its other components and they all play an integral role in language generation.

This is another key component of the Minimalist Program. Each aspect of the computational

system that has been retained is absolutely necessary and described in minimal terms.

---

[48] A very technical argument for this point may be found in Berwick and Chomsky *Why Only Us: Language and Evolution* Ch. 2
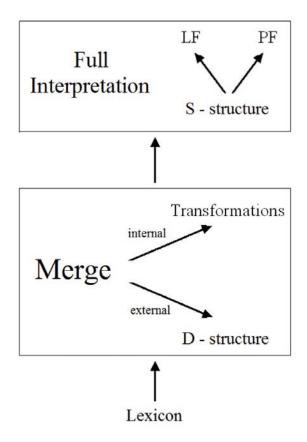
Figure 1.1: A mental map of the Linguistic computational system from Mukherji, 2014 p. 325

The lexicon can be described metaphorically as a mental dictionary.[49] It stores "word-like" units with definable properties that act as the fundamental material upon which Merge operates. The term "word-like atoms" is generally used in place of words to reflect the notion that even single words may actually be a combination of more fundamental lexical items. For example in English, "played" is a single word. Linguists view this term as the product of two more fundamental lexical units that have already been combined via Merge. "Play" is viewed as a lexical item with distinct properties such as meaning, pronunciation, syntactic category

---

[49] The discussion of the lexicon is derived from Carnie, Andrew. Syntax: A Generative Introduction Oxford: Blackwell Publishing, 2002. Print p. 165-186.

Michael Anthony Ciaramella
The Convergence of Schenkerian Music Theory and Generative Linguistics

(verb phrase, noun phrase etc.), and other lexical properties like the theta criterion (See Carnie,

2002 for a more detailed explanation of this). Linguists also view the "-ed" as the second lexical

item active in the word "played". This too has lexical content detailing the tense of phrase. In

other languages, this can even be reflected as a separate phonological unit that may appear in a

different portion of the phrase in relation to the verb. Whether existing as a single

phonological unit as in "played" or separated within a given phrase, linguists view this word as a

generation from more basic lexical content. Therefore, the lexicon is a finite set of word-like

units, like "-ed" or "play" with distinct properties relating to their semantic, syntactic, and

phonological detail.[50]

Infinite linguistic creativity arises when the Merge acts upon the content of the

lexicon.[51] As mentioned earlier, Merge is a binary set theoretic operation that combines lexical

items into higher-order phrases. It always acts on two single units and combines them into a

single, larger unit. Returning to the example of the English word "played," Merge acted upon

the two units "play" and "ed" and combined them into the single, larger unit, "played."

Formally, the possibilities of Merge operations can therefore be defined as follows:

(1) A + B → (AB)   (combination of two lexical items)

(2) C + (AB) → [(AB)C]   (combination of a lexical item and a larger Merge product)

(3) (AB) + (CD) → [(AB)(CD)]   (combination of two Merge products or phrases)[52]

---

[50] Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 317-319 gives a concise explanation of the interaction between Merge and the Lexicon.
[51] The discussion from this paragraph through page 37 of this thesis regarding Merge as a binary set theoretic operation working externally or internally stems from Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 315-326, though situated within a larger discussion of the history of generative linguistics in p. 243-326.
[52] See ibid. 318

where single letters represent lexical items such as "play," and two letter combinations

represent single units representative of previous Merge operations (termed phrases). Items

grouped in parentheses or brackets represent single phrases generated from previous Merge

operations. In each case, Merge only combines two single units into a larger unit. Importantly,

however, Merge cannot actually act upon a phrase itself as a single unit. Instead, some lexical

content from within the phrase, known as the "head" of the phrase, projects itself as the

definable property of that phrase. In essence, the most important information governing the

content of a phrase projects as the head of the phrase. This is the mechanism by which we

obtain verb phrases and noun phrases, among others. Therefore, we can refine options (2) and

(3) from above to say that in the cases of a phrase Merging with a lexical item, or a phrase

Merging with a phrase, it is actually true that the head of the phrase is subject to the Merge

operation and that the other content of its associated phrase (i.e. its constituents) are carried

along with the head into the new Merge product. In this way, we can think of all Merge

operations as a case of possibility (1) from above, whether A and B be fundamental lexical items

themselves or heads of phrases. This is because a head is really a projection of lexical content

from within the phrase. Therefore, what appears to be a higher-level combination is always

representative of multiple consecutive binary Merge operations involving fundamental lexical

content.

Since Merge executes binary combinations, the issue of unit order immediately arises.

This can be illustrated formally with the with the simple example below:

(4) C + (AB) $\rightarrow$ [C(AB)]

(5) C + (AB) $\rightarrow$ [(AB)C]

(6) C + (AB) → [A(C)B]

In the simple case of a lexical item Merging with a phrase (or rather the head of a phrase), three

ordering options become apparent. Either the lexical item may be ordered before the phrase

as in (4), the lexical item may be ordered after the phrase as in (5), or the lexical item may

insert itself as a unique unit somewhere within the phrase as in (6). All three options are

possible within the theoretical context of the linguistic computational system. In practice

however, a given language only chooses one of the three orderings consistently. This ordering

is known as the "parameters" of a given language. The field of generative linguistics

hypothesizes that when a child is acquiring language early in life, she is attuned to the syntactic

ordering of different types of phrases.[53,54] Upon acquiring a given language, the child develops

a sense of the order in which Merge must combine units in order for the product to be

intelligible within her language. As will be discussed shortly, this does not mean that Merge will

not combine units in other orders as well, but that in order for the product to pass through the

Logical interface prior to an utterance, it must present itself with the proper ordering.

Earlier paradigms of generative linguistics developed complicated rules for the order in

which units may be combined to meet grammatical requirements in a given language. These

rules became increasingly complicated trying to account for syntax in each of the world's

languages. This is partially to blame for why generative linguistics was called into question as a

field in the early 1980's. The Minimalist Program does away with the complicated rules of

syntactic generation in preference for the simple and elegant Merge mechanism, which can

---

[53] Carnie, *Syntax: A Generative Introduction* p. 13, which also suggests that language is an
innate instinct within all of humans, rather than something solely learned by external stimuli.
[54] Berwick and Chomsky *Why Only Us: Language and Evolution* Ch. 1 also discusses this process.

manifest in one of two forms: internal Merge or external Merge. External Merge is the most fundamental operation and characterizes syntactic possibilities 1-5 from the preceding examples. For this to occur, one unit of lexical information is simply combined linearly with the next unit, either before or after it. External Merge does not inherently impose an order upon the linear combinations. Rather, these arise within a sociocultural specific context as a language's parameters. The specific orderings of a given language may arise from any number of reasons from cultural or societal norms to sensorimotor limitations and are not of particular focus for the Minimalist Program. The important fact is that external Merge combines units A and B into either (AB) or (BA). This generates what linguists call the D-structure of a linguistic utterance. All utterances in all languages have a particular D-structure that results from the external Merge operations requisite for a given structure.

External Merge, however, does not account for theoretical possibility (6) which results in the structure [A(C)B]. This is instead the product of internal Merge. Internal Merge is the process by which an external Merge product Merges again with itself. Therefore, internal Merge essentially rearranges the order of constituents within a phrase without changing the lexical content of the overall unit. In order for internal Merge to occur, external Merge must occur first, generating a D-structure. Then, internal Merge may cause a constituent within the phrase may be reintegrated into the phrase in a different place. This results in what linguists call the S-structure of an utterance. This used to be short for "surface structure," just as D-structure is short for "deep structure." Linguists now resist the terms "surface" and "deep" because it reflects an older paradigm in generative linguistics in which it was hypothesized that the mechanisms by which the D-structure and S-structure were created were carried out

separately within the mind and in unique ways. Now, however, it is known that both S-structure and D-structure are a product of a single mechanism, Merge. Linguists now believe that the only difference between the two is that S-structure is representative of the actual ordering of words heard in a sentence after all internal Merge operations have been performed and that D-structure is representative of the same sentence before internal Merge. One may also ask why internal Merge operations are necessary at all for a linguistic utterance and why D-structures are insufficient. The answer to this question relates to the phonetic level of representation through which all Merge products must pass before realization. Transformations by internal Merge occur to meet phonological requirements of a language and to streamline the phonological baggage associated with a sentence. Logically, however, a D-structure may be entirely interpretable semantically, but unpronounceable given the morphophonological requirements of the human vocal apparatus, or hands in the case of sign language. Therefore, internal Merge operations alter the phonological content, but not the semantic content of a phrase. As an extension, internal Merge operations are also important for the expressive and aesthetic qualities of language, though these are also particularly culturally defined.

The discussion of internal and external Merge brings to light two of the most important definable properties of human language that set it apart from other animal communication systems. These are the notions of hierarchy and recursion in syntactic grammar.[55] Hierarchy refers to the idea that consecutive strings of words do not necessarily need to result from the same Merge operation. Rather, constituents may present with any number of words between

---

[55] Derived from a personal communication with Dr. Mukherji

them that may relate at other, higher levels of syntax but not in the immediate Merge

operation. This enhances the generative potential of human language by allowing ideas to be

embedded within others. Therefore, in a hierarchical grammar, smaller ideas, or modifying

ideas, may be embedded within an overarching idea. This powerful aspect of human language

allows for precision and complexity of expression and communication. These expressive and

communicative qualities are unique to human language, therefore making it a far more

advanced communication system than that of other animals.[56] Additionally, the idea of internal

Merge allows for a phenomenon called recursion, in which a syntactic structure is cyclically

embedded within itself, generating a seemingly infinite linguistic structure. For example,

consider the sentence "Enzo played ball". The phrase "the baker" may be embedded within

this sentence via an external Merge operation followed by an internal Merge operation,

resulting in the sentence, "Enzo, the baker, played ball. A recursive operation would embed

another descriptive category of the noun, Enzo, to the sentence in the same. For example, one

might say, "Enzo, the baker, who is tall, and a friend of Jane, played ball. In this sentence, three

phrases of similar syntactic categories (all modifying or adding to information about Enzo) were

embedded within a single sentence. Formally, one could conceive of repeating this process

infinitely and therefore generating an infinite linguistic structure. This is proof that at least in

theory, human language holds infinite possibilities and therefore an infinite potential for

creativity.[57] This is important to keep in mind whether studying language or music from a

scientific perspective such as this one. We must take care to showcase the generative capacity

---

[56] Berwick and Chomsky *Why Only Us: Language and Evolution* Ch. 2 covers this topic in detail.
[57] See Carnie, *Syntax: A Generative Introduction* p. 13-14 for a succinct discussion of recursion.

of these uniquely human phenomena and treat them in a way that highlights their creative and expressive potential, rather than attempt to reduce them.

As mentioned previously, within the Minimalist Program, the sole characteristic of successful Merge operations is that they meet the requirements of both the logical and phonetic interfaces within the mind. This is another important facet of the Minimalist approach because it again streamlines the mental map of the computational system to solely the interaction of one computational mechanism with two separate mental interfaces. If the product passes through both of these interfaces successfully, the phrase is said to have achieved what linguists call "full interpretation." This means that the phrase is both intelligible and pronounceable (this includes other sensorimotor operations such as the movements of the hands during sign language).[58] It is important to notice that these are the bare minimum requirements for any linguistic utterance and that the Minimalist Program says nothing about how meeting these requirements may be achieved. In other words, this approach says nothing about the actual qualities of language that express things like emotion, aesthetic impact upon others, or even how it generally sounds. These are left to other sociocultural or anatomical factors that lie outside the scope of this paradigm.[59] The Minimalist Program is only concerned with the pragmatics of generating structures that can be conveyed to others successfully with a meaning that is at least interpretable.

I will now briefly summarize the general aspects of the linguistic logical and phonetic forms. A more in-depth explanation can be found in Berwick and Chomsky's *Why Only Us:*

---

[58] Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 323-324
[59] Ibid. 326

Language and Evolution, which also places their explanations in the context of evolutionary

feasibility.[60],[61] This aids in understanding how and why these structures may have developed

within the human mind. The logical form of language is mediated by a conceptual-intentional

interface within the mind and ensures that the forthcoming phrase is interpretable. The issue

of interpretability versus interpretation is important to linguistics and centers around the

logical interface. To address how the logical interface targets interpretable structures, without

automatically applying a specific interpretation, we must consider the minimum requirements

for an intelligible phrase. This can be tricky if one considers the fact that a phrase might in fact

be interpretable without an individual being able to interpret it, since interpretation is

dependent on other cognitive factors such as intelligence. Therefore, for linguists to determine

if a phrase is interpretable, while bypassing the necessity of their own interpretation, it is

important to develop a formal definition of interpretability. In a formal sense, interpretability

requires that each aspect of lexical content from the resulting Merge operations agrees with

the others. This accounts for things like subject-predicate coupling, tense agreements,

modifiers, and other aspects such as these. This does not necessarily entail that the phrase

must be perfect grammatically in the sense we learned throughout grammar school, since it is

known that most people speak intelligibly in so-called ungrammatical, or dialectic language.

Thus, grammar in the common sense of the world has some sociocultural influence. The formal

definition of a logical phrase must therefore correspond to a deeper level of linguistic

---

[60] The discussion from page 39 to the top of page 41 in this thesis derives partially from Berwick and Chomsky *Why Only Us: Language and Evolution* Ch. 2 and the reference in footnote 61.
[61] Also see Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 315-326, 450-451
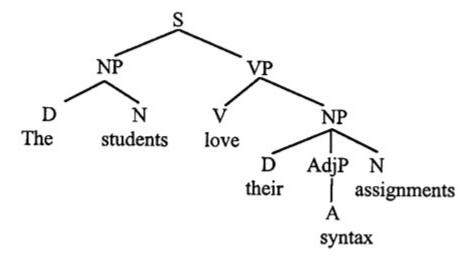
representation, relating to the agreement of actual lexical content. This is a complicated aspect of the Minimalist Program and one in which this paragraph only scratches the surface. However, for the purposes of applying this methodology to music, recognition that the logical interface works with a formal, mathematical definition of interpretability relating to lexical content retained through Merge operations will suffice.

The phonetic form of language is mediated by the sensorimotor system within the brain, whose sole purpose is to generate phrases that are pronounceable in some way. An interesting discussion of the sensorimotor apparatus for language and its evolution may also be found in Berwick and Chomsky's *Why Only Us: Language and Evolution.* The basis of phonetic form in language is that the sensorimotor apparatus imposes restrictions on the kinds of phrases that are capable of pronunciation. For example, the vocal apparatus is restricted to only a set range of frequencies and only a certain degree of precision for pronunciation. Therefore, Merge outputs that fall outside these constraints are incapable of conveying through the sensorimotor apparatus and therefore cannot be realized as language. Interestingly, Berwick and Chomsky claim that language was first acquired within the mind as a mechanism for thought (a controversial claim indeed). Therefore, Chomsky claims that Merge is not restricted by phonetic form for thought, but is restricted when we try to convey thought to others in the form of language. There is also a rich field devoted to the morphophonological constraints on language leading to the sounds (or sights) we know, but this is also outside the scope of this thesis. For this purpose, it is only important to know that Merge is limited externally by phonetic representation and that language is only realized when Merge products successfully pass through both their logical and phonetic restrictions.

In the interest of making this discussion slightly more tangible, I will briefly present a simple example of a linguistic tree to explain how a linguist might interpret the sentence, "The students love their syntax assignments," (taken from Carne (2002) pg. 32).[62,63] Linguists use syntax trees to denote the hierarchical structure of a sentence and represent the operations that took place prior to its final construction. An example from Carnie, 2002 shows a tree of the example sentence, though from a slightly earlier paradigm in linguistics. Still, it serves its purpose here as an example of a linguistic syntax tree and the accompanying methodology.

Figure 1.2: A syntax tree of the sentence, "The students love their syntax assignments." (from Carnie, 2002, pg. 32).

Within this tree, constituents are represented by lines that extend vertically to the same level. For example, "the students" are shown as constituents within the noun phrase (NP) portion of this sentence. Every juncture represents the head characteristic of the constituents, and therefore which unit projected its lexical content to a higher level. Returning to the NP, "The

---

[62] Sentence example and tree from Carnie, *Syntax: A Generative introduction* p. 32. The Merge accompanying this tree on p. 42-43 is my own since this tree is reflective of a slightly earlier paradigm in generative linguistics.
[63] Additional examples of Merge analyses of linguistic trees may be found in Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 243-326

students," it is apparent that the word students (or more precisely, "student" modified in number by "s") projected its lexical content as a noun to a higher level of the tree. Importantly, one will notice that every juncture only ever has two lines extending downward from it. This is representative of the fact the Merge always operates in a binary fashion. Therefore, whether Merging with the head of another part of the tree, or with a lexical item itself, each new juncture only results from two units. Examining the verb phrase (VP) juncture, the word "love" Merged with the head of phrase "their syntax assignments," which is the noun, "assignment." In order to create the phrase, "their syntax assignments," four Merge operations must occur. First, "assignment" must be modified by with a lexical item by external Merge, resulting in, "assignments." Next, "assignments" must undergo external Merge with the modifier (denoted D), "their," indicating possession. Finally, the word "syntax" must Merge with this phrase, first externally, before internal Merge places it between the previous two constituents. Therefore, the larger VP higher in the tree is representative of lexical content indicating that some people are in possession of syntax assignments (the NP) and that they love those assignments (the verb). A final external Merge operation takes place linking the head of the NP to the head of the VP, resulting in the final sentence. Before this sentence was realized, the logical interface would ensure that all lexical properties necessary for interpretability matched from the given Merge operations, and that the necessary transformations occurred for phonological articulation by the sensorimotor pathway.

Section C: Schenkerian Music Theory and the Minimalist Program

Dr. Mukherji's effort to connect Schenkerian music theory and the Minimalist Program is part of a broader claim that music and language share a common computational system within the human mind. Therefore, attempting this connection depends on the existence of several common structures between music and language. First, there must be a lexicon of musical units analogous to the linguistic lexicon of word-like units. Additionally, there must be evidence of a hierarchical transformational grammar within music, similar to that of a syntactic tree, and that this grammar must show the same binary linear combinations as executed by Merge. Finally, there must exist some kind of logical interface and phonetic interface that limits musical structures to those that are intelligible and capable of articulation. Mukherji claims that arguments for these connections lie within Schenkerian music theory, which I will detail now.[64]

The first, and perhaps most controversial claim necessary for this connection, is the existence of a musical lexicon containing a finite set of musical units with definable properties. An argument for the musical lexicon originates from Schenker's notion of *stufe,* or "scale-step". Schenker argued each area of a musical passage can be characterized by specific scale degrees corresponding to their fundamental harmonic structure. Therefore, in the case of a musical passage beginning with a tonic in C-major, Schenker would argue that we should view this beginning as an elaboration of the 1,3, and 5 scale-steps. Additionally, a set of harmonies that

---

[64] The remaining content of this chapter expounds Dr. Mukherji's connection between Schenkerian music theory and the Minimalist Program and will therefore only be cited once here. The discussion can be found in Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 328-451. Additionally, aspects of this connection were clarified via personal communication with Dr. Mukherji.

prolongs a certain passage, such as a plagal cadence in the context of the final tonic of a piece, can be viewed as a *stufengang*, or collection of pitches centered around a single scale-step. Therefore, Schenker used the background, middleground, and foreground levels of his analysis method to indicate the primary scale-steps populating a musical passage, in order of importance. These scale-steps generated structural motives at each level corresponding to what we generally think of as a harmonic structure. The *Ursatz* within the background, which for him was always I-V-I, corresponded to the most fundamental scale-steps of the piece's background structure. Middleground elaborations were therefore manifest of subordinate scale-steps to those of the background and served to construct *stufengang* structures about the background scale-steps. Finally, the foreground served to further elaborate the *stufengang* developed through the background and middleground structures, employing voice-leading operations, ornamentations, and even more subordinate scale-steps. Importantly, Schenker believed that fundamental content from the background scale-step projected itself at each level of a piece in ways that elaborated and expanded upon it. He termed this phenomenon *Asckomponierung.*

Mukherji believes Schenker's idea of a hierarchical organization of scale-steps from the background to the foreground actually represents the beginning of a generative theory of tonal music. The idea that a scale-step has definable properties and may be organized in a hierarchical arrangement parallels the idea of word-like units within a lexicon. Further, the idea that a scale-step may project itself at every level of structure, as those of the *Ursatz* do, also parallels the idea that lexical content projects as the head of a Merge operation. Therefore, combinations of scale-steps in which one scale-step projects its information as hierarchically

more important from the background to the foreground can be seen to parallel the idea that

when two lexical items are Merged together, one item projects its information to higher levels

of syntactic representation and defines the content of a phrase. For this claim to hold weight,

however, it needs refinement to apply to all musical idioms since as it stands, it would appear

that the notion of the scale-step would only apply to polyphonic idioms with harmony similar to

that of the Western Classical idiom.

Schenker's definition of the *Urlinie* substantiates the argument that the universal lexical

basis for music might be found in his theory of the scale-step. Mukherji claims that Schenkerian

theory is primarily a melodic one, as opposed to a harmonic one. This originates from the idea

that the *Urlinie* is constructed from scale-step tones. Therefore, in the case of the 3-2-1 *Urlinie*,

scale degree 3 would come from the initial tonic scale-step, scale degree 2 would come from

the dominant, and scale degree 1 would come from the final tonic. If this is the case, then a

melody should already hold the scale-step content it is generated from and therefore hold the

lexical content inherent to scale-steps. The purpose of harmony in this line of thinking would

be to externalize the harmony already present in a melody. This is commonly seen in

monophonic idioms that often arpeggiate or unfold a choral harmony linearly, therefore serving

as evidence that a melody may contain inherent scale-step content. Conversely, music in lead

sheet notation may be assumed to contain a *Urlinie,* even if not outwardly notated, on the basis

that it has a logical harmonic progression. This link allows for an extension of Schenkerian

theory to potentially all genres of music, since it applies to both monophonic and polyphonic

textures. Therefore, Mukherji's conclusion is that there exists a set of chord-like structures,

based in Schenker's notion of the scale-step, that serve as the basis for all of the world's music

and that these are the fundamental content upon which Merge operates within the mind's computational system.

The only issue remaining is Schenker's sentiment that the basis of music's lexical content must be the projection of a I-V-I *Ursatz* into a 3-2-1 *Urlinie.* Given the ideological limitations that restricted Schenker to the so-called Germanic master composers, this might be true for music as he understood it. However, in today's global society, it would evade our sensibilities to believe that this truly could be the basis for all music. However, if we view the *Urlinie* and the *Ursatz* as culture-specific artifacts of a more basic computational mechanism, then we can overcome this limitation. Mukherji claims that the *Ursatz* and *Urlinie* arise from Merge operations of basic scale-step content, the structure of which has become part of the logical form of Western Classical music, thus explaining its prominence. There is nothing intrinsic to this progression, however, that would lend itself to any kind of universality. Therefore, we would expect that all musical idioms have some form of an *Urlinie* and an *Ursatz*, based in the idea of the scale-step, but that each idiom's structure may not bear similarity to any other. Even the scales used to derive these culturally defined progressions may be different from idiom to idiom. The only necessary commonality is the existence of a hierarchical and recursive grammar centering around a scale-step based lexicon upon which Merge operates. Further, as an aspect of human biology, it would also be expected that the products of Merge be subject to the logical and phonetic interfaces limiting their possibilities.

In music, the issue of interpretability versus interpretation brought about by the logical interface is perhaps even more pertinent than in language. Meaning in language may be complex, but its nature lends itself well to a single interpretation. For example, all readers

would likely understand the sentence, "The Godfather Part II is my favorite movie," in the same way. Because of this, deducing what might be the formal, logical basis for interpretability seems manageable with regards to language because we can come to consensus regarding what a meaning is in the first place. This allows a platform upon which we can define the notion of interpretability formally in terms of lexical, grammatical agreement. Music, however, does not share this characteristic with language. Meaning within music is hotly debated from all corners of the music academy without even the possibility of consensus. Some even claim that music itself is meaningless. Among those who acknowledge meaning exists in music, there exists an infinitely many possible interpretations of a piece that can be defended as valid.

Given this inherent plurality, the notion of interpretability versus interpretation becomes an ever more important aspect of the Minimalist Program and one that sets it apart from other analytical methodologies. The logical interface through which Merge outputs are processed before finalization ensures that the lexical content from each scale-step within the structure align formally in a way that is interpretable, but does not actually impose an interpretation to the music. Interpretation, therefore, is not the subject of the mental computational system generating music and something that is applied through other sociocultural forces.   Even information like harmonic function is said to arise semantically after the syntactic Merger is processed through the logical interface. Mukherji explains the concept of "harmonic function" as a cultural artifact of the Western Classical idiom that arose as a means of semantically describing Merged syntactic structures within the idiom, but claims that they arise after the fact rather than as a necessary component of the computational system.

Exploring the characteristics of the universal musical lexicon and its relation to logical

form is an area in need of further study to substantiate the connection between Schenkerian

music theory and the Minimalist Program. Mukherji presents one hypothesized aspect of the

musical lexicon as being a scale-step feature relating to the circle of fifths. He terms this

characteristic the "cf-feature." A cf-feature is a numerical label applied to each pitch in relation

to its position on the circle of fifths. The actual value of this numerical label is arbitrary, but the

difference between cf-values of two consecutive or Merged scale-steps is of importance. Figure

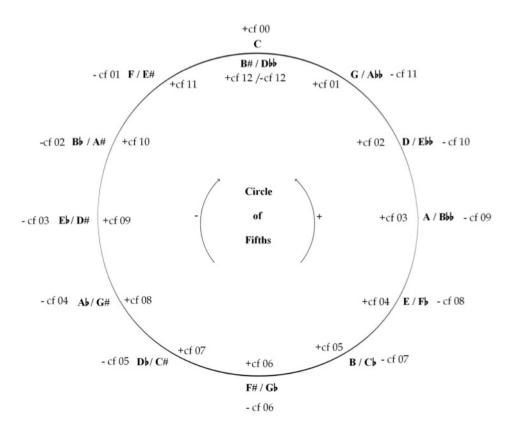1.3 shows a map of the cf-values for each area of the circle of fifths.

Figure 1.3: cf-feature labels for the circle of fifths from Mukherji, 2014 p. 338.

The cf-feature provides a starting point for the formal logical definition of an interpretable tonal

Merge structure and reflects the parameterization of Merge ordering found in language. For

the Classical idiom, it is conventional to move in a counter-clockwise direction around the circle

of fifths. This is made most clear by the prominence of the authentic V-I cadence ending

virtually every classical piece. Formally, Mukherji suggests that the parameter of Merge

ordering within the Western Classical idiom requires that the Merging of two scale-steps follow

the cf difference, Δcf=01. This is reflective of a change between adjacent positions on the circle

of fifths moving in the counter-clockwise direction. For example, take an authentic cadence

between D^7 and G major. This would correspond to the change cf=02 to cf=01. The difference

between these (02-01) is Δcf=01. Therefore, the Merger would be syntactically valid and order

linearly with the D^7 chord moving to G major. Additionally, the Merger of Δcf=00 is always

possible in any idiom and would result in no chord change at all. This seems trivial, but is

actually necessary for a logically interpretable phrase so that it begins and ends on the tonic.

This is the basic idea behind all scale-step changes parameterized within the Western Classical

idiom.

From this basic discussion, questions immediately arise regarding how a deceptive

cadence or a cadence involving a diminished seven chord can be accommodated within this

structure. Mukherji claims that there exists a second scale-step lexical relationship between

scale degrees related by a third. He calls this a specific "pitch space," just as the circle of fifths is

a pitch space of fifth intervals. In thirds-based space, pitches separated by a third can be seen

to essentially hold the same cf-value. Evidence for this characteristic is the existence of the

relative minor, in which a scale with the same number of accidentals dictates both major and

minor depending on the starting note. The order of this operation, however, should be

parameterized just as fifths-based space is parameterized to move counter-clockwise in the

Western Classical idiom. Mukherji observes that in Western Classical music, thirds-space is parameterized in moving down a third, as opposed to up. Therefore, C major can be seen to hold the same cf-value as A minor, but not E minor. From this, it follows that cf movement may proceed covertly between scale-steps related by a third in order to fulfill the syntactic rule that cf-values move counter-clockwise in fifths-based space. For example, in the case of a V-vi deceptive cadence. It would actually be the case that the initial Merger was between V and I, therefore fulfilling the required $\Delta cf=01$, and then covert movement between I and its third-space relation vi generated the articulated vi chord. In this case, the I chord is present grammatically and syntactically, but not actually articulated in the structure. This also explains a progression such as (vii^o/V)-V-I. While the (vii^o/V)-V seems to break the syntactic Merger rule by resulting in $\Delta cf=+05$, the structure may be explained by using a covert third-space relation between (vii^o/V) and (V/V). In this case, the syntactic Merger was between (V/V) and V, which would generate the proper $\Delta cf=+01$, and covert movement occurred within third space moving (V/V) to (vii^o/V).

The notions of fifths-based cf-values and thirds-based space allows us to explain other canonical structures within the Western Classical idiom in addition to other Western idioms such as the blues. Take the progression I-IV-V-I, which is the basis for a conventional Western Classical phrase and the origin of Riemannian harmonic areas. Using fifths-based cf-values and third-based space, we can develop the following syntactic tree for this generic progression shown in Figure 1.4.
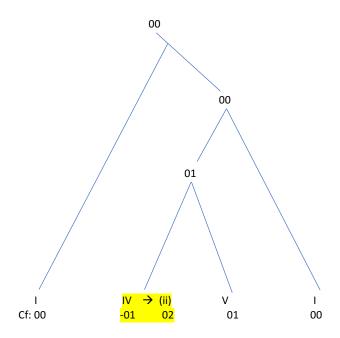
Figure 1.4: A syntactic tree showing cf-values and Mergers for the I-IV-V-I Progression.

In the above syntactic tree, lexical items originating from a leftward-up slanting arrow are

subordinate to the other lexical item they Merge with, which originates from a rightward-down

slanting line. Therefore, leftward-branching units function to prolong the hierarchically

superior unit. Each juncture represents a Merge operation in which either the heads of

previous operations or a lexical items are Merged. The resulting phrase links the initial tonic to

the final tonic, resulting in an overall phrase cf-value of 00. This is a requirement of Western

Classical logical form, that the overall phrase is characteristically Δcf=00. We can now walk

through each Merge operation that resulted in the case of a cf=00 Merging with a cf=00 to

generate the overall phrase. First, the scale-step of hierarchically least importance is the IV

chord. This cannot immediately Merge with the V chord because it breaks the cf-value syntax

rule by creating Δcf=+02. Theoretically, it could Merge with the initial tonic, creating Δcf=+01,

but this would forbid subsequent Merging with the V chord. Instead, covert third-based

movement occurs between IV and ii, which the lexical content of the ii projects. This is noted

within the yellow box of Figure 1.4, which takes the cf-value of 02. The notion of third-based

movement between IV and ii also seems plausible in historical context, since in the Western

Classical idiom, IV is often articulated before ii in a pre-dominant context. Following third-

based movement to the cf-value of ii, this scale-step may Merge with the V scale-step, again

generating the required $\Delta cf=+01$. In each case, the latter scale-step projects its lexical

information up the tree, reflective of the parameterization of Western Classical music. This

branch can now Merge properly with the final tonic, again creating the desired $\Delta cf=+01$ and

projecting the tonic's cf information to the highest level of the tree. The final Merger between

the tonic cf head of the previous Merger and the initial tonic may now occur, create the

permissible $\Delta cf=00$ which is also required for a logically interpretable phase. As we can see, the

cf-value argument between fifths-based space and thirds-based space can explain this

canonical phrase within the Western Classical idiom, reflecting the counter-clockwise

parameterization of the idiom for logically interpretable phrases.

As mentioned earlier, the cf-feature of lexical scale-steps should not be taken as the sole

lexical characteristic of musical units. Rather, it is likely just one factor among many others,

though more work needs to be done to elucidate these. For example, there must be some

modification to indicate tonicizations versus full modulations. Additionally, music that exists in

other modes may feature other lexical content, or observe a different parameterization of the

cf-feature. For example, Mukherji worked with blues and blues-influenced rock idioms and

revealed that the parameterization of these idioms is exactly the reverse of the Western

Classical idiom. In these genres, scale-steps proceed clockwise about the cf-modified circle of fifths and therefore have the *Ursatz* of I-IV-I and a cf structure of Δcf=-01. In this argument, a V chord seen at the end of a phrase would represent a transformation of the D-structure I-V-IV-I, as seen in the final bars of the blues. Therefore, the above argument should not be taken as final, but rather as the beginning of a more rigorous characterization of the musical lexicon. Still, this will suffice for the case studies that will be presented in Chapter two, since they are all from Western derived idioms.

The final aspect of the musical computational system to describe is the phonetic interface through which music travels prior to realization. This is the most complex and heterogeneous aspect of the computational system discussed thus far and I will only cover the major points from Mukherji's discussion here. He first notices that most theorists treat tonal and rhythmic aspects of music as separate and distinct entities. Therefore, musical rhythm is often viewed from an externalist perspective. One major facet of the Minimalist Program, however, is that it cohesively accounts for phonology within language from an internalist perspective, taking it to be inseparable from the other aspects of language. In order for the link between music and language within the Minimalist Program to be substantiated, an internalist stance toward musical rhythm must also be taken. Mukherji therefore argues that pitch has an asymmetric influence on rhythm, but that there is an interplay between the two that generate an overall musical structure.

This is done first through the observation that when two chords are presented, rhythm is inevitable introduced to music. This is because the chords must be ordered in some way, therefore separating their articulation in time. This is the definition of rhythm. He then

distinguishes between metered rhythm, which has a regular pulse, and rhythm in general, which only requires the ordering of notes. Metered rhythm can be seen as a special case of the more general rhythm and one whose characteristics are defined by an interplay between formal syntactic factors and extra musical factors. For example, the Chopin mazurkas purposefully incorporated rhythms known from Polish dance. Therefore, one may expect to find an explanation for the specific accents within the footsteps of the Polish dances, since there are anatomical limitations to how the human body can dance. This is an extra-musical factor of musical rhythm.

The discernment of internal musical factors affecting rhythm are more relevant for the characterization of the musical computational system. To deduce what these might be, Mukherji first observes that music often has a regular distribution of harmonic structures in relation to rhythm. Specifically, harmonic units generally appear in divisions of two or four. Mukherji believes that this is not a coincidence or due to external factors. The regular division of harmonic units may be the result of the binary Merge operations. If all rhythmic phenomena are driven by Merge's binary operation, then at an abstract level all rhythm must be binary. This is complicated by the Merging of a head to another head or another lexical item, but the basic premise holds that the origins of rhythm must be binary. Therefore, Mukherji goes on to argue that this binary operation manifests in harmonic structures of two or four units. Further, claims that what might appear to be a deviation of this kind of division is often resolved at a deeper level of structure. Therefore, a D-structure may present with the proper binary arrangement and by transformations be converted to an S-structure with a seemingly ungrammatical rhythmic structure. Further, voice leading operations and prolongations may

alter rhythm at the surface, but not at deeper levels. This interaction also works in the other direction. Rhythmic irregularities at the level of D-structure may be resolved by insertion of another beat in to the S-structure, satisfying voice leading or other phonological concerns. Mukherji elaborates on this idea to a much greater extent, but the general idea follows that pitch asymmetrically influences rhythm through Merge and that D-structures and S-structures interact to reinforce rhythmic structure.

Conclusion

This chapter developed not only a theoretical basis for a connection between music and language within the Minimalist Program of generative linguistics, but also a methodology rooted in Merge analyses of musical structures. The goal of such analysis is to characterize the process by which the mind generates musical structures with infinite creativity from a finite set of musical inputs. The next chapter will be devoted to applying this theoretical and methodological framework to three original case studies and investigating whether it can accommodate musical structures from variety of musical idioms and time periods.

Chapter 2: Case Studies from Western Idioms

In this chapter, I present three case studies which investigate aspects of the musical computational system. The first case study will be the opening phrase of the first movement of Mozart's K. 332 piano sonata in F Major, which fits squarely into the Western Classical idiom. The next two case studies are "Georgia on My Mind" and "Blue Bossa" and fall under the general category of Jazz, though their influences are vastly different within the idiom. My goal for presenting these analyses is threefold. First, each piece showcases different aspects of generative musical grammar and will be used to demonstrate the ability of the methodology presented in Chapter One to accommodate a diversity of musical structures. Second, I will argue for a common grammatical structure underlying each of these pieces and present this as evidence for the validity of the Minimalist Program-Schenkerian music theory connection. Finally, I will expound a theory that the difference between cyclic and linear song forms lies within the characterization of the final Merge operation in the piece's grammatical structure and will use Mozart K. 332 as an example of a linear structure and "Georgia on My Mind" and Blue Bossa" as examples of two cyclic forms with different grammatical underpinnings.

The case studies presented here were chosen because they are each derived from a Western idiom, but not the same idiom. For a work of the magnitude of a thesis, as opposed to a dissertation, it is useful to choose pieces that broadly fall into a similar musical paradigm. This saves the additional work of translating musical structures and terminology from their original idiom into another in which there might not be an exact translation. By this, I am not suggesting that Western Classical music and Jazz share any particularly strong sociocultural

connection, but rather that they simply share a similar set of terms by which we describe the music (harmony, melody, etc.) It may be there is some cultural continuity between the idioms and this would perhaps explain some of their descriptive similarities in addition to their musical similarities, but that is beyond the scope of this thesis. This does not discredit cross-cultural investigation within the generative framework. Rather, these studies are essential to proving the psychological basis for the linguistic/musical computational system. This is only to suggest that such cross-cultural studies often require the workload of a thesis in and of themselves, if not a dissertation. However, the chosen pieces do allow for strong cross-idiom comparison of musical grammar, which is another key facet of proving the psychological basis of this generative theory. The works presented here span almost two centuries and originated on different continents. Commonalities of musical syntax would be expected across this timespan and geographic span if they were reflective of a psychological reality, whereas there might not be commonalities if there was simply some loose cultural continuity between eighteenth century Western Europe and twentieth century United States. Therefore, the cross-idiomatic similarities I will present in what follows will be taken as evidence for an underlying computational similarity.

Section A: Mozart Piano Sonata K. 332 Mvt. 1

The first movement of Mozart's K. 332 F major piano sonata is written in conventional sonata-allegro form.[65] Theme A from the exposition, which is the focus of this analysis, spans mm. 1-12. The theme is presented in F major and consists of three shorter four measure

---

[65] Score used in this analysis was obtained from Mozart, Wolfgang Amadeus. *Sonata No. 12 in F Major* K. 525 Bryn Mawr, PA. Theodore Presser Company, 1960.

phrases. Phrase one spans mm. 1-4 and arpeggiates an F major triad from F4 to F5 and features a pedal F as the lowest note of the bass arpeggiation. This pedal will be significant later for arguing that this entire phrase prolongs the initial tonic. The next phrase spans mm. 5-8 and prolongs the tonic in F major. The melody sequences downward from F5 to F4 while outlining an F major triad. The final phrase spans mm. 9-12 and features a descending scalar passage outlining the predominant harmony before an authentic cadence between mm.11-12 with C major resolving to F major.

An investigation of the generative process behind Mozart's construction of the opening phrase depends heavily upon the lens we choose for analysis. For example, if we take a broad lens and view the phrase in context with the entire Exposition, we may see the entire twelve measures as a prolongation of the initial tonic and a single scale degree. Alternatively, we may take a smaller scope and arrive at a different conception of the phrase. For the present investigation, we will take mm. 1-12 as a single unit that comprises a coherent musical structure in and of itself. In other words, we will treat it as a linguist would treat a sentence within a larger paragraph or essay. Within this phrase, Mozart's mind would have employed the Merge mechanism and acted upon the specific triadic lexical content that is common to the Western Classical idiom to arrive at an interpretable phrase. Our task is to reveal the logical structure inherent within the phrase and visualize the specific Merge operations requisite for its construction.

If we were to perform a Schenkerian analysis on just the opening phrase, we are immediately confronted with a difficult choice for discerning the *Urlinie*. The initial five
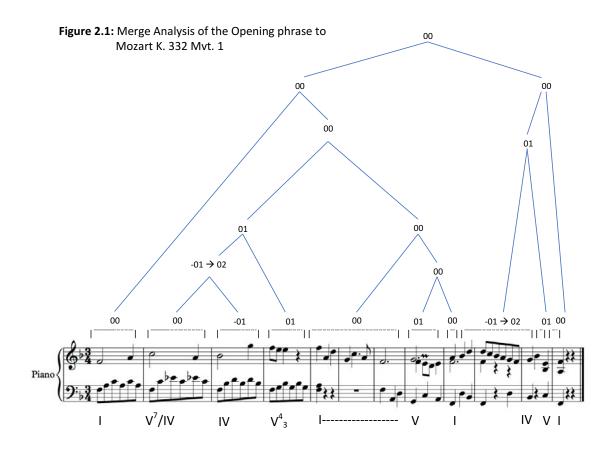
measures outline an F major triad from F4 to F5. The final seven measures then descend from F5 back to F4. Therefore, one could argue that this phrase is comprised of an 8-7-6-5-4-3-2-1 *Urlinie* in which the *Kopften* (highest note of the *Urlinie*) is revealed following the initial ascent in m. 5. This would require viewing the initial five measures as a prolongation of the tonic and an elaboration of scale degree one. Following the descent, the dominant chord on the third beat of m. 11, which precedes the final tonic in m. 12, would then be viewed as the structural dominant of the *Ursatz* in which scale degree two would be implied and replaced with scale degree seven by a voice leading operation. This analysis, however, leaves questions regarding the integrity of the 8-7-6-5-4-3-2-1 *Urlinie*. Schenker seems to generally reserve this *Urlinie* for larger structures.[66] It is therefore unlikely that the entire scale could descend structurally within the span of seven measures without a sequence or some other measure to harmonically justify the descent. Additionally, scale degree seven is omitted from the descent. At its best, this phrase would contain a quick and incomplete *Urlinie*, if we were to view it this way.

Alternatively, we could take scale degree five on the first beat of measure two as the *Kopften* and say that the phrase has a 5-4-3-2-1 *Urlinie*. This may seem more plausible, given that it is revealed more quickly durationally and unfolds over the course of the entire phrase, rather than the last seven measures. If this were the case, then we might argue that the descent from scale degree five to scale degree two is quicker, spanning mm. 2-8. In this analysis, scale degree two would begin structurally on the downbeat of m. 8 in conjunction with the dominant harmony. Mm. 9-11 would then prolong scale degree two by elaborating the dominant with a short phrase before the final resolution in mm. 11-12. This would also reveal

---

[66] This idea came about through a discussion with Professor Shirish Korde.

the structural dominant of the *Ursatz* as the C major seven chord on the third beat of m. 11.

There are also reasons to question this analysis. First, this would require the initial statement of the *Kopften* to occur over a V^7/IV chord. Intuitively, this weakens the statement of the fifth scale degree because it is associated with motion to the IV chord, rather than the tonic. Since Schenker believed the *Urlinie* to derive from structural tones within the *Ursatz*, we would expect the statement of scale degree five to coincide with the tonic, rather than this pre-dominant sonority. Additionally, the only statement of scale degree four within this analysis follows the *Kopften* immediately in m. 3. Since IV is the predominant sonority to which the secondary dominant in m. 2 is leading, this would require us to acknowledge that two structural tones of the *Urlinie* would unfold within one phrase area and over the course of two measures. Again, this seems unlikely.

A Merge analysis of the opening phrase assists in clarifying the proper designation of the *Ursatz* and *Urlinie* within the opening phrase. Therefore, I will first detail the phrase's Merge structure to reveal an underlying logical cohesion rooted in the cf-feature, before explaining the proper Schenkerian diagram of the phrase. Figure 2.1 shows a Merge diagram of the opening phrase. Since this phrase is derived from the traditional Western Classical idiom, we would expect a parameterization of the cf-feature in which scale-steps proceed counter-clockwise around the circle of fifths, as demonstrated in Chapter One through the analysis of the canonical I-IV-V-I chord progression. Indeed, this is what we find within this phrase. The entire structure is constructed of three smaller phrases, each of which embodying some realization of the I-IV-V-I progression.

Figure 2.1: Merge Analysis of the Opening phrase to Mozart K. 332 Mvt. 1

The first step in deriving this phrase's Merge structure is to identify the major structural components that comprise the head of each section. To begin with the initial tonic, we find that the F major chord articulated in m. 9 is the head of the initial tonic phrase. Therefore, we can view mm. 1-9 as a leftward prolongation of the initial tonic in F major.[67] For this argument, the V chord in m. 8 must first Merge with the tonic in m. 9, which creates Δcf= 01. If we allow the tonic cf=00 content to project as the head of the phrase, then the head may Merge with the tonic elaboration that appears in mm. 6-8, creating the allowable Δcf=00 Merger. The overall projection from mm. 6-9 would therefore be the tonic cf=00 content. Before Merging m. 6-9

---

[67] This means that the structurally prolonged tone appears after its prolongation (See Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 363-365).

with the dominant, however, we must deal with the V^7/IV and IV sonorities in mm. 2-3. To account for this, we must acknowledge that the F dominant seven chord moving to the B flat major chord represents allowable counter-clockwise motion around the circle of fifths. Therefore, the chords in mm. 2-3 may Merge to generate Δcf=01 and the lexical content of IV, which is cf=-01, projects as the head of the phrase. This creates the first dilemma for the Merge structure, in which the cf=-01 cannot Merge with the dominant, since it would create Δcf=02, and it cannot Merge with the tonic in m. 1, because we would be unable to join this with the rest of the phrase. We can circumvent this dilemma by employing thirds-space movement between IV and ii following the initial Merger of V^7/IV and IV. Now, the lexical content of ii would define the head of this phrase, which is cf=02. This can undergo an allowable Merge with the dominant in m. 4, in which the dominant's lexical content projects as the head of the phrase. This can then participate in a Merger with the tonic phrase from mm. 6-9, in which the tonic's lexical content projects to the higher order phrase. Finally, this phrase Merges with the initial tonic chord in m. 1, creating a logical phrase of Δcf=00. As we can see, the cf-feature of the entirety of mm. 1-9 is actually the tonic's lexical content, indicating that the initial tonic is prolonged over the course of the first nine measures.

To generate the remainder of this phrase, we must take mm. 1-9 as a prolongation of the initial tonic with cf=00 lexical content. What remains is an articulation of the canonical I-IV-V-I progression that was detailed in Chapter 1. First, the IV chord in m. 9-10 undergoes thirds-space movement to ii and takes the cf value of 02. This undergoes an allowable Merger with the dominant in m. 11, in which the cf=01 lexical content project as the head. Next, the cf=01 head merges with the final tonic in m. 12 and generates a cf=00 head. Finally, this phrase can

Merge with the tonic prolongation in mm. 1-9, generating the logical Δcf=00 and an interpretable twelve measure phrase.

This Merge analysis helps to reveal the true *Urlinie* and *Ursatz* that defines the first phrase.[68] Shown in Figure 2.2, the opening phrase of the movement is best characterized as a statement of a 3-2-1 *Urlinie*. We can understand the entirety of mm. 1-9 as a prolongation of the initial tonic, in which the tonic's structural tones are arpeggiated upward to F5 in a projection from *Ursatz*, and then downward to F4. Finally in m. 9, the *Kopften* is stated on the downbeat as scale degree three. Therefore, we can view the first nine measures as a leftward prolongation of the *Kopften* leading up to its articulation in m. 9. This analysis is corroborated by the pedal F underneath the arpeggiated harmony in mm, 1-4, which clues us into what might be Mozart's intuition about the function of the first four measures. Further, the two dominant

Figure 2.2: Schenkerian Diagram of Mozart K. 332 Mvt. 1 Opening Phrase

---

[68] Several conversations with Professor Shirish Korde were also helpful for discerning this phrase's *Urlinie.*

sonorities are situated with scale degree two in the bass, indicating a weak authentic cadence. These arise from voice leading considerations and elaborate the tonic further. The remaining mm. 9-12 correspond to the descent of the *Urlinie* in a 3-2-1 overarching structure. The second tone of the *Urlinie,* scale degree two, is articulated at the beginning of measure 11, just before the structural dominant of the *Ursatz* is articulated on the third beat of that measure. Within this analysis, the descending scalar passage in m. 10 is taken to outline the predominant sonority, which is corroborated by its role as a cf=02 harmony in the Merge analysis. A voice leading operation replaces scale degree two with the leading tone when the structural dominant is articulated. The final tonic is then articulated on the downbeat of m. 12.

There is one remaining issue to discuss with respect to this phrase. The issue regards the question of why the V^7/IV chord in m. 2 is treated as an elaboration of the predominant sonority that Merges with the adjacent IV chord, while its cf-value is the same as the initial tonic. There are two approaches to understanding this secondary dominant chord grammatically. First, we can treat the V^7/IV chord as a new lexical item, owing to the addition of the flat seven scale degree. By saying that the addition of this scale degree adds lexical content to the chord, then we must include it as a distinct unit grammatically. This is the approach taken in the above analysis. The chord is allowed to retain the same cf-feature since there is still a fifths relationship between the secondary dominant and the IV chord, but its content Merges in a fifths-space relation with the IV chord in m. 3, rather than elaborating the tonic in m. 1.

Alternatively, we may say that our label of the F dominant seven chord as a V^7/IV is

actually a semantic label applied to the chord in retrospect and extra-grammatically.[69] Within

this argument, we assume that since the initial tonic in m. 1 and the chord in m. 2 hold three

notes in common, that their lexical content is the same. Therefore, we can view mm. 1-2 as a

single elaboration of the initial tonic, rather than as two distinct entities. This would still allow

for an interpretable generative structure, but would require that mm. 1-2 are Merged together

and at a higher level of structure, after the IV-V-I have already Merged in mm. 3-5.

This argument is important for the next two case studies in this chapter, since Jazz

harmonies often contain many scale degrees outside of the basic triad. Therefore, there must

be some way of distinguishing where the threshold lies for a cultural-semantic interpretation of

additional chord tones and when lexical content actually changes. Intuitively, it seems likely

that the elaboration of 9ths, 11ths, 13ths, etc. in Jazz do not constitute new lexical content,

since performers add or remove these tones relatively frequently. Alternatively, there seems to

be something special about scale degrees three and seven in music that add lexical content.

These are considered the guide tones of Jazz and create a dominant seven chord in Western

Classical. Therefore, since the tone added to m. 2 of the opening Mozart phrase was a minor

seventh above the root, it was treated as though it contributed to the lexical content of the

chord. Decorations of other tones will be taken to hold semantic content, but not affect the

grammatical interpretability of the structure. That said, a more rigorous investigation of the

effect that scale degree additions have on chordal lexical content is needed. Such an

---

[69] This argument relates to Dr. Mukherji's discussion of interpretability vs. interpretation and an example of a time in which it is hard to discern a personal interpretation from an aspect of grammar that aids with interpretability. See Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 357 for an explanation of harmonic function and Logical Form.

investigation would proceed by analyzing all the circumstances in which the tonal alteration

occurs and observing whether patterns exist in the chords preceding and following the altered

chord. If the chord seems to function differently when the altered tone is present in

comparison to when it is not, then there may be a grammatical and lexical effect from the

altered tone.[70]

Section B: Georgia on My Mind

"Georgia on My Mind" is a jazz tune originally written by Stuart Gorrell and Hoagy

Carmichael in 1930.[71] Since then, it has been covered by a number of artists, perhaps most

notably Ray Charles in 1960. The song is quite versatile in terms of instrumentation and can be

performed in every context from big band to a singer/keyboard combo. Overall, the song has

an A-B-A form, where A is the setting of the chorus and B is a bridge in the relative minor. Due

to its versatility and improvisatory quality, this analysis will focus around the cf-Merge structure

rather than an *Urlinie*. The music is scored as a lead sheet with a melody and chord symbols.
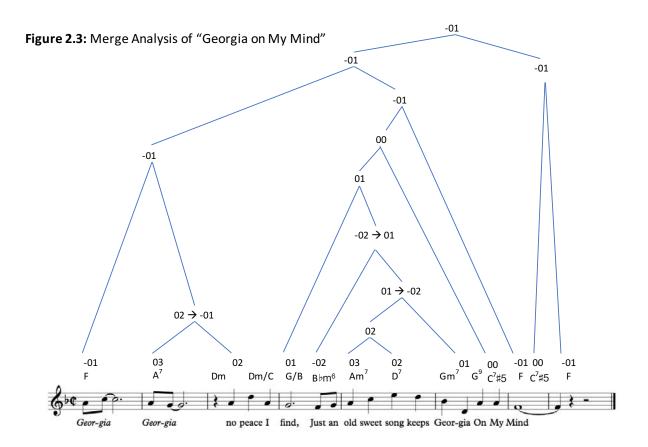
Since performers are left to realize the texture through their means, the *Urlinie* may appear

within the texture and not necessarily the melody. Without such codification, we can only use

the Merge structure to discern the logical basis of the song. However, since the *Urlinie* is

manifest from the Merge structure, if we can discern a logical basis rooted in Merge for the

---

[70] Mukherji's discussion of morphophenomic rules in language was an important part of the thinking here. (Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 256) These are rules relating to the pronunciation of words and the constraints on the vocal apparatus when certain sounds are present. Later, Mukherji, p. 357-362 make an interesting comparison of these rules to voice leading and chord alterations.

[71] The lead sheet for this analysis was obtained from Gorrell, Stuart and Carmichael, Hoagy. *Georgia on My Mind* Milwaukee, Wi. Hall Leonard Corp, 1998.

song, we may also assume that an *Urlinie* exists within the piece, though not actually notated in

the score.

The Merge analysis shown in Figure 2.3 reveals two complete phrases within the final

verse of the song. The first phrase is articulated in mm. 1-7, and the second phrase proceeds in

mm. 7-8. It may seem redundant to denote mm. 7-8 as a phrase on its own, but lexical content

is introduced within this measure through the C dominant seven chord. Our reluctance to

denote this as a phrase stems from extra-grammatical consideration of its semantic importance

in the content of the song. Since no new lyrical content is introduced and the melody is static,

we semantically interpret mm. 7-8 as the ending of the initial phrase. Grammatically, however,

the introduction of lexical content requires that there be a formal explanation for the final two

measures. The formal explanation for mm. 7-8 is also quite simple. It takes the form of a I-V-I

Figure 2.3: Merge Analysis of "Georgia on My Mind"

chord progression in which the initial tonic is the head of the first phrase. This provides

additional evidence for our intuition that the phrase is of lesser importance than the initial

phrase, since it is grammatically simple, but does not reduce the fact that it is a grammatically

logical phrase.

The initial phrase in mm. 1-7 is considerably more complex and interesting than the final

phrase. Overall, the phrase follows the canonical I-IV-V-I model that is also common to classical

music. The way the phrase unfolds, however, is considerably more complex. The initial tonic is

stated in mm. 1-3 with the lyrics, "Georgia, Georgia, no piece I." The A^7 chord in m. 2 serves as

a secondary dominant to the Dm chord in m. 3, which is the relative minor of F major. Mm. 2-3

participate in a fifths-space Merge in which the lexical content of Dm projects as the head.

Already, thirds-space movement is required to reveal the cf-feature of F major, which is also the

key of the piece. Following the thirds-space movement, mm. 2-3 can Merge with the initial

tonic for a Δcf=00. Cf=-01 from the F major chord projects as the head of this phrase area.

Measures 4-5 and the first three beats of m. 6 constitute what would be semantically

interpreted as the pre-dominant phrase area. We must keep in mind, however, that semantic

interpretation and the labeling of phrase areas occurs after an interpretable Merge structure

already exists. Therefore, it is more precise to treat mm. 4-8 as an overall unit which prolongs

the final tonic and articulates the structural dominant. To build this phrase, the Am^7 on beats

one and two of m. 5 and the D^7 on beats three and four of m. 5 participate in a fifths-space

Merge generating Δcf=01. The lexical content of D^7 projects as the head of this phrase. The

Gm^7 and G^7 in m. 6 are treated formally with the same cf-feature and Merge with the head of

the previous Am^7-D^7 Merger to generate another Δcf=01. The lexical content of the Gm^7

projects as the head of this Merger and undergoes a thirds-space movement to obtain the

lexical content of B flat. The head of this phrase can then Merge with the B flat in m. 4 and

project cf=-02 to the higher-order phrase. The thirds-space movement that occurred in the

previous Merge is then reversed to re-articulate the lexical content of G major, which can then

Merge with the initial first inversion G chord found in m. 4. Overall, this phrase projects the

lexical content of cf=01 to its head and holds the semantic interpretation of a pre-dominant

prolongation.

The initial phrase is generated from three Merge operations following the creation of

the pre-dominant area. First, the pre-dominant constituent from m. 4 to the first three beats of

m. 6 Merges with the $C^7\sharp5$ chord in the last beat of m. 6. Lexical information from the $C^7\sharp5$

chord projects as the structural dominant to the higher-order phrase. As seen in previous

examples, the head of the structural dominant product Merges with the final tonic in m. 7,

indicating that at the highest level, mm. 4-7 serve as a prolongation of the final tonic through

the pre-dominant and dominant phrase areas. Finally, the first phrase is generated from the

Δcf=00 merger between the initial tonic head and the final tonic head.   The overall structure is

then generated from another Δcf=00 Merger between the first and second phrases.

Two issues must be addressed regarding the logical structure of "Georgia on My Mind."

First, there seems to be a disconnect between the lyrical and grammatical content within the

song in m. 4. The lyrics end the phrase, "No piece I find," following a descending line. At the

same time, however, the structure reveals pre-predominant motion toward the dominant and

the overall final tonic. This calls into question the role of the G/B chord in m. 4. We can

reconcile this apparent contradiction by acknowledging that the grammatical interpretability of

a logical phrase says nothing about its actual interpretation. Therefore, one answer to this contradiction is to say that the lyrics do not matter in this instance and that our anxiety toward the phrase stems from extra-grammatical, semantic interpretation of the phrase as serving a pre-dominant function. With the pre-dominant label, we are primed to think of the phrase as signaling motion to the dominant and therefore are made uneasy by the seemingly conclusive lyrics. However, if we resist applying the conventional pre-dominant label to the chord until after we discern the phrase's logical structure, then this contradiction disappears.[72]

If this conclusion seems a bit unsatisfactory, it might also be useful to speculate as to how one might keep the pre-dominant label and account for the lyrical shape. Let us accept the pre-dominant label for mm. 3-6 and interpret it as moving toward the dominant and later resolution to the tonic. With this label, we must feel the G/B as the beginning of motion toward the structural dominant, rather than the end of a phrase. The lyrical content of, "No peace I find," becomes at issue again. If we look with a slightly larger lens, however, we will see that the lyrics move quickly from the F4 on the word, "find" to an eighth note pickup leading into the next measure. Additionally, this pick up is followed by continuous quarter note motion until the final tonic resolution. We may analyze this motion as a signification that the lyricist also interpreted the G/B chord as the beginning of pre-dominant motion and felt the need to move away from the F4 on "Find," with a pickup and sustain motion until the final tonic. This may have been used to reflect the lyricist's anxiety expressed within the words, "No peace I find," and later resolution by the comforting idea of his home in Georgia. Therefore, we may

---

[72] Again, see Mukjerji, "Generative Musical Grammar-A Minimalist Approach" p. 357 for the discussion of harmonic function and Logical Form.

account for this contradiction by saying that it was a conscious choice by the lyricist in response to a specific interpretation of the phrase.

The second issue with the above grammatical structure is the bidirectional thirds-space movement found between the B ♭ and G cf-values. Within the Western Classical idiom, thirds-space movement is unidirectional depending on tonality. Major keys move down to relative minor and minor keys move up to relative major. It would be unconventional, however, for a major piece to move from the tonic up to the key of scale degree three. This is at issue within the "Georgia on My Mind," since the logical basis of the final phrase depends on both directions of thirds-space movement. A satisfactory answer to this dilemma would require more extensive analysis of the genre. Nevertheless, I will speculate as to why this may be. It is possible that thirds-space is always bidirectional and that the appearance of strict unidirectional movement might arise from semantics and culture. This would mean that the motion from IV to ii in Western Classical music would arise from extra-grammatical consideration and that either direction is actually plausible. This would explain the bi-directionality of "Georgia on My Mind." Alternatively, the sense of key in Jazz music might be more fluid than in Western Classical. Therefore, the appearance of bidirectional movement in Jazz actually reflects rapid movement between major and the parallel minor. This theory may be reinforced by the many accidentals and additional chord tones layered upon basic triadic structures. Therefore, the same parameterization may exist in Jazz as in Western Classical and the appearance to the contrary would be the result of tonal fluidity in Jazz.

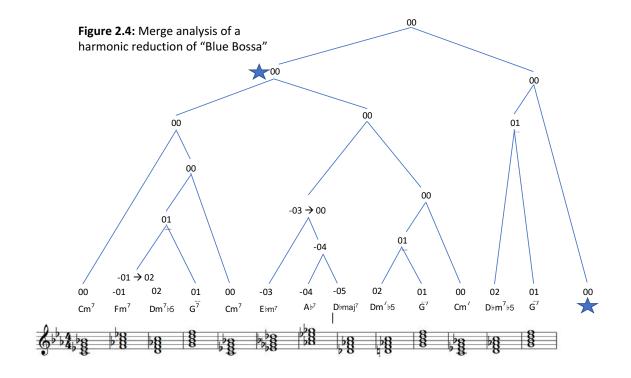## Section C: "Blue Bossa" and Cyclic Forms

"Blue Bossa" is an instrumental jazz song written by Kenny Dorham in 1963. This song was chosen as a case study because it exemplifies one type of a cyclic song structure.[73] "Blue Bossa" is written in a simple sixteen bar form in the key of C minor. Its form lends the song well to improvisatory solos which may proceed for any number of iterations of the song's form. I will argue that the grammatical structure of "Blue Bossa" establishes it as a continuous cycle by requiring the highest level Merge operation to occur within a single phrase, rather than two distinct phrases. In doing so, the logical coherence of the song becomes dependent upon it being a cycle. This contrasts with the typical definition of a logical musical structure, in which Merge events occur first between separate phrases and then with itself through an internal Merge, if applicable. Just as in "Georgia on My Mind," this analysis will chiefly concern the harmonic structure of the song, due to the improvisatory quality of the music. Since the *Urlinie* is conceived from structural tones, the existence of a logical structure within "Blue Bossa" will be taken as evidence for an underlying *Urlinie* as well, even if it is articulated on the lead sheet. Additionally, the analysis to follow will utilize a harmonic reduction of the piece in order to fit the line onto a single page for analysis.

The parameterization of "Blue Bossa" is identical to the previous two case studies within this chapter. Structural tones proceed counter-clockwise around the circle of fifths, revealing a i–ii-V-I underlying progression. Allowing for thirds-space movement, we can consider ii within this structure as identical grammatically to IV in the previous case studies. Whereas Mozart K.

---

[73] The lead sheet obtained for this analysis was obtained from Durham, Kenny. *Blue Bossa* Petaluma, CA. Sher Music co., 1988.

332 and "Georgia on My Mind" articulated the IV sonority and underwent covert movement to

ii for Merging with the dominant, this piece proceeds straight to the ii chord.

Figure 2.4 reveals three underlying phrases within "Blue Bossa." Each chord is

articulated for either one or two full measures in common time and is the overall harmonic

structure is set to a descending sequential melody from G5 to G4. Beginning again with the

initial tonic, the F minor and D minor seven flat five chords in mm. 2-3 participate in thirds-

space movement to project the single cf=02 lexical content into the first Merger with the G

Figure 2.4: Merge analysis of a harmonic reduction of "Blue Bossa"

major seven in m. 4. Lexical information from the G major seven projects as the head of this

Merger and constitutes the structural dominant of this phrase. Just as in the previous case

studies, the dominant phrase Merges with the C minor chord in m. 5 to project the tonic's

information to the higher-level phrase. Finally, the initial phrase is completed by a Δcf=00

Merger with the initial C minor chord in m. 1 and the C minor head of the Merger of mm. 2-4.

A choice must be made regarding the three chord segment following the first phrase

consisting of E flat minor seven, A flat dominant seven, and D flat major seven. With a cf-value

of 05, the E flat chord cannot Merge with the initial phrase on its own. Subsequently, the D flat

chord is a minor second from the following D half diminished seven chord and therefore cannot

Merge with that either. This three-chord sequence appears to disrupt the structure of what

would be two simple i-ii-V-I progressions. A potential solution to this dilemma is to treat these

three chords as a prolongation of the initial tonic of the second phrase by employing thirds-

space movement. The A flat chord can Merge within conventional fifths-base space with the D

flat chord and projects lexical information as cf=-04 from A flat. This product can participate in

another fifths-space with the E flat chord. The resulting three-chord constituent therefore

prolongs the E flat chord through a descending fifths sequence and constitutes a leftward

prolongation of the chord. At this point, we may employ thirds-space movement to reveal the

constituent's status as the initial tonic of the second phrase. E flat therefore undergoes third-

space movement to its relative minor, and key of the piece, C minor. This alleviates the

dilemma presented earlier and allows these three chords to Merge into the second phrase with

the following D minor, G dominant, and C minor seven chords, as has been seen many times

now within this thesis. Additionally, the first two phrases can Merge into one larger tonic

phrase with the Δcf=00 and overall 00 lexical projection.

There is one question that must be addressed before the remainder of this structure is

detailed. This regards the adjacent D flat major seven and D half diminished seven chords.

Since these chords share the same upper three notes and only differ by a half-step in the root, one could argue that the lexical content of these two chords is identical. Therefore, we might treat these chords as a single unit and have them Merge with the adjacent dominant. We cannot do this, however, because it generates a logical problem with the preceding E flat and A flat chords in which there is seemingly no place for them within the structure. Hence, the choice was made to treat this three-chord sequence as a leftward prolongation of E flat major, revealing a coherent phrase. Still, one could inquire as to why two chords that share the same third and seventh can be treated as grammatically separate units, even when they are adjacent to each other in the musical texture. It also seems possible that the motion of a half step arises from some voice leading consideration or perhaps an aesthetic choice.[74] The issue is really another manifestation of the previous question of which chord tones constitute lexical content and which do not. In the Mozart K. 332 example, the minor seventh above F major in m. 2 altered the tonic's lexical content to function as a pre-dominant texture and Merge with the IV chord. Alternatively, "Georgia on My Mind" and "Blue Bossa" are riddled with seventh chords and other decorations and yet, these do not change lexical content. Further investigation is required to elucidate the contents of the musical lexicon and its sensitivity to alterations.

Analysis of the final phrase in "Blue Bossa" reveals the logical construction of a cyclic form. The score of this piece ends on the dominant chord in C minor and a G in the melody as well. Without resolution to the tonic, it appears as if the entire song's structure is made illogical, since it ends with a cf of 01, instead of 00, and the Merge structure is incomplete. This

---

[74] Again, the discussion of morphophenomic rules might explain this phenomenon. (Mukherji, "Generative Musical Grammar-A Minimalist Approach" p. 256, p. 357-362)

goes against intuition, however, since "Blue Bossa" has attained the status of a Jazz standard and is performed regularly across the world. It would seem unlikely that such a popular song would be logically uninterpretable. Additionally, "Blue Bossa" is not alone in this structure. The score of "Georgia on My Mind" contains two endings, one that signifies a return to the top of the form, and one that signifies the end of the song. The first ending also shows a dominant chord at the end of the structure. More generally, the common "turn around" found at the end of the Blues structure and many other Jazz standards ends with some form of the dominant. Given its ubiquitous status within these idioms, it seems likely that there is a grammatical basis for ending structures that lead back into the song and those that signify closure.

The logical basis for cyclic song forms relies upon the argument that a single phrase may participate in more than one Merger. In a linear structure, such as the Mozart K. 332 example, each Merge consists of two distinct phrases. Closure within a linear musical structure is signaled by the final remaining musical unit Merging into the final phrase. In a cyclic form, however, chords must be repeated more than once. We have two options for explaining this phenomenon. The first explanation is that the repetition of a musical form is external to its grammar. A phrase in language may be repeated endlessly without a grammatical explanation for its repetition. Therefore, it is possible cyclic forms are just linear forms that a performer or composer chooses to repeat. This seems likely for instances such as the repetition of a sonata's Exposition. The Merge structure comes to a close at the end of the exposition and the performer is free to choose whether to repeat the section or move on to the Development.

This solution fails to explain songs that do not reach a logical conclusion at the end of the song's form, as seen in "Blue Bossa" and the more general "turn around" structure.

Therefore, a second explanation for cyclic song forms is that a Merge operation connects the end of the form to the beginning. This solution was shown in Figure 2.4 using a star. In this case, the G dominant seven chord can Merge in fifths-space with the initial tonic phrase which includes the first eleven chords. Therefore, this large phrase serves as the initial tonic within the form and also as the final tonic of the overall structure. The overall Δcf=00 is therefore constructed by the tonic phrase Merging with itself and is denoted through a star on Figure 2.4, since tree diagrams generally proceed linearly, rather than cyclically. This explanation entails that the logic coherence of "Blue Bossa," and more general "turn around" structures is dependent upon the existence of the cycle. In cases such as these, the repetition of the musical form is not a mere performance choice, but a necessity and part of the piece's compositional genesis. In order to break the cycle, performers must replace the tonic phrase from the initial eleven chords with another tonic to complete the song's logical structure. This may be completed through the simple addition of a tonic chord to the end of the song form. The statement of another tonic chord adds an additional musical unit to the texture and allows the preceding dominant to Merge with it, rather than to cycle.

Conclusion

The case studies presented in this chapter were from compositions spanning almost two centuries. Still, striking commonalities within their underlying logical structure were revealed through Merge analysis. Each piece demonstrated counter-clockwise cf parameterization and the canonical I-IV-V-I or I-ii-V-I progression that is indicative of this motion. Additionally, each piece utilized thirds-space movement to resolve apparent contradictions at multiple levels within its logical structure. Questions were raised, however, regarding whether thirds-space

movement is parameterized in a unidirectional manner, or if bi-directional movement is always allowed. Additionally, a grammatical distinction was drawn between linear musical structures, such as the Mozart K. 332 Mvt. 1 example, and cyclic structures like "Blue Bossa." Within a linear structure, each Merge operation is populated by distinct lexical content and the final structure is signified both by the $\Delta cf=00$ Merger and the exhaustion of all available lexical content. In cyclic forms, the final Merger must occur either with the head of a previous phrase, or with a chord that is not present in the score.

Several questions were raised throughout this chapter regarding aspects of the musical lexicon. Characterization of the musical lexicon is essential to this theory, since it affects our grammatical treatment of a chord. For example, the introduction of an E flat in m. 2 of Mozart K. 332 was taken as an important event from a lexical content, shifting the chord's function from tonic to pre-dominant. Alternatively, the many seven chords in "Georgia on My Mind" and "Blue Bossa" were taken as decorations of the basic triadic structure, though this was called into question by the motion from D flat major seven to D half diminished seven. Future work in this field should revolve around a more thorough interrogation of the musical lexicon. One way to do this would be to examine non-Western idioms in the manner presented above and perform a cross-cultural analysis of structural characteristics and function in a Merge diagram. Without any cultural or instrumentation similarity between the idioms, we gain clearer insight into the purely musical factors that drive Merge. For example, each of the case studies presented here featured the basic triadic structure as the commonality between each lexical item. Perhaps this is not the case for all idioms. A cross-cultural investigation of musical syntax would aid in elucidating what the bare minimum criteria for musical lexical content are.

A more elaborate characterization of the culture-specific issues of semantics and patterns in

lexical content may then proceed from the thorough definition of the musical lexicon.

Chapter 3: An Exploration of Musical Syntax

Program Note

The study of music places us in a paradox to which there is seemingly no escape. It is both fully an art and a science and we are challenged with reconciling these two seemingly opposing universes when we enter into any scholarly investigation of music. This string quartet is my attempt to reconcile the artistic and scientific sides of music. I've spent my senior year completing a thesis entitled "The Convergence of Schenkerian Music Theory and Generative Linguistics: and Analysis and Composition," in which I explore a recent psychological theory connecting music theory to work done in the field of generative linguistics. The end goal of such a pursuit is to develop a theory of cognition which accurately characterizes the mental process by which music is generated. Much of my year has been spent reading through contemporary advances in music theory, music psychology, neuroscience, and linguistics relating to the questions of how music and language can both allow for infinite creativity of expression, and yet have a definable and regular form. Throughout the year, I have engaged music as a scientific object, taking recordings and scores to be my data. All the while, however, I also wanted to keep in mind that music is, at its core, an art form. This is where the string quartet fits in. I've allowed the idea of music as a science to incorporate itself in my creative process. There are aspects of this piece that creatively represent ideas from psychology and linguistics. Others draw upon claims about linguistic and musical syntax as their structural and compositional foundation. Overall, my goal was to complement my study of music as a science with a work of art and bring a point of connection between these seemingly opposing universes.

Michael Anthony Ciaramella
The Convergence of Schenkerian Music Theory and Generative Linguistics

The first movement of this piece is entitled "Hello." This movement engages the curious moment in which a young child acquires language. For the first several months of life, babies are bombarded with what appear to be foreign sounds. Somehow, without being told what is noise and what is stimulus, babies discern on their own what language sounds like. They begin to babble, exploring the capabilities of their vocal apparatus. Soon, they utter their first word, likely repeating a sound they have heard from a parent of loved one. Before long, they will do something remarkable. They will utter a phrase that has never been said before. This creative utterance represents a significant step in the development of language, since it diverges from a simple repetition of sound and into the infinitely creative capacity of language. The field of generative linguistics claims that this creative utterance, and the infinitely many that will follow throughout the child's life, is an indication that humans are disposed to be linguistic. It claims that humans are attuned to linguistic syntax and acquire certain parameters and sound patterns that constitute a specific language upon an innate knowledge of linguistic form. Thus, the first utterance of a creative remark is evidence that the child has sufficiently acquired a kind of library of sounds and structures which their mind's linguistic system can work with.

In "Hello," I have represented the process by which we acquire language. The movement is defined by a single melody. It is fragmented, altered and at some points unrecognizable. A competing texture masks and reveals it, sometimes hiding it entirely. Yet, the rhythm attempts to push forward, bringing ever more clarity to the melody as the texture becomes more complex. Eventually, the movement reaches its climax, in which the melody and the surrounding texture each appear separate and distinguishable from the other. The melody, like that first creative utterance, is then heard in its entirety.

The second movement is entitled "Forms." This movement draws upon the technical

characterization of the linguistic computational system put forth by the Minimalist Program in

generative linguistics, which has recently been proposed to serve as the fundamental linkage

between music and language in a cognitive sense. The Minimalist Program postulates that

there exists a finite computational system within the human mind that is fit with the inputs of

linguistic structures (word-like units with definable properties) and generates outputs by a

simple binary combination called "Merge". By this mechanism, Merge combines two word-like

units to form a single, larger unit. These binary combinations are performed at higher and

higher levels, leading to theoretically infinite linguistic structures. These structures contain a

hierarchy of linguistic information relating to the level at which they were integrated by Merge.

There are only two constraints placed on the Merge output before language is realized. First, it

must pass through a sensorimotor interface, which mandates that the output must be capable

of articulation in some way, whether this be through the vocal apparatus, sign language, or

another medium. Second, it must pass through a logical interface, which mandates that the

product is intelligible (meaning that it is capable of comprehension, not particularly that any

one person comprehends it).

In "Forms," I have worked some of the structural characterizations of Merge and the

linguistic computational system into my compositional process. First, there is no order imposed

on the binary combinations executed by Merge. Thus, different languages combine inputs in

different orders. This accounts for one aspect of how there can be seemingly thousands of

unique languages in the world if they all originate from a single computational system. In

"Forms", I take a single rhythmic pattern or melody, break it into smaller parts, and reassemble

them into a seemingly new larger structure. Sometimes this entails incorporating a retrograde of a fragment, or extending a passage by applying additional linear combinations. Since Merge imposes no order on its outputs, the beginning of the melody of rhythmic passage is not always represented at the beginning of the new structure. Additionally, Merge does not demand that binary combinations be in close proximity to each other. They may be split by another binary combination, or even part of another combination by a linguistic transformation. Thus, I embed parts of structures into others and develop larger structures with smaller structures embedded within. Finally, the theoretical capability of the linguistic computational system is infinite. It is only limited by the constraints of our memory's capacity, brain's energy, and otherwise our lifetime. In "Forms," I shrink the meter over the course of the piece from quadruple to double by eighth note increments. I then return to the quadruple meter without retracing through the intermediate steps. This represents the idea that a linguistic structure can be infinitely long and complex, or as simple as a single binary combination.

Bibliography

Alamia, Andrea, Oleg Solopchuk, Alessandro D'Ausilio, Violette Van Bever, Luciano Fadiga,

Etienne Olivier, and Alexandre Zénon. "Disruption of broca's area alters higher-order

chunking processing during perceptual sequence learning." *Journal of cognitive

neuroscience* (2016).

Agawu, Kofi "How We Got Out of Analysis, and How to Get Back In" *Music Analysis,* 2004: 267-

286.

Berwick, Robert C., and Noam Chomsky. Why Only Us: Language and Evolution. Cambridge,

MA: MIT Press, 2016.

Bharucha, J., and C. L. Krumhansl. "The Representation of Harmonic Structure in Music:

Hierarchies of Stability as a Function of Context." Cognition 13.1 (1983): 63-102.

Blacking, John. How Musical Is Man? Seattle: University of Washington Press, 1973.

Carnie, Andrew. Syntax: A Generative Introduction Oxford: Blackwell Publishing, 2002. Print.

Durham, Kenny. *Blue Bossa* Petaluma, CA. Sher Music co., 1988.

Gjerdingen, Robert O. "Gebrauchs-Formulas." Music Theory Spectrum 33.2 (2011): 191-99.

Gorrell, Stuart and Carmichael, Hoagy. *Georgia on My Mind* Milwaukee, Wi. Hall Leonard Corp,

1998.

Herdener, Marcus, Thierry Humbel, Fabrizio Esposito, Benedikt Habermeyer, Katja Cattapan-

Ludewig, and Erich Seifritz. "Jazz drummers recruit language-specific areas for the

processing of rhythmic structure." *Cerebral Cortex* 24, no. 3 (2014): 836-843.

Koelsch, Stefan. "Significance of Broca's area and ventral premotor cortex for music-syntactic

processing." *Cortex* 42, no. 4 (2006): 518-520.

Koelsch, Stefan, Tobias Grossmann, Thomas C. Gunter, Anja Hahne, Erich Schröger, and Angela D. Friederici. "Children processing music: electric brain responses reveal musical competence and gender differences." *Journal of Cognitive Neuroscience* 15, no. 5 (2003): 683-693.

Koelsch, Stefan, Sebastian Jentschke, Daniela Sammler, and Daniel Mietchen. "Untangling syntactic and sensory processing: An ERP study of music perception." *Psychophysiology* 44, no. 3 (2007): 476-490.

Krumhansl, Carol Lynne. "Statistics, Structure, and Style in Music." *Music Perception: An Interdisciplinary Journal* 33, no. 1 (2015): 20-31.

Kunert, Richard, Roel M. Willems, Daniel Casasanto, Aniruddh D. Patel, and Peter Hagoort. "Music and language syntax interact in Broca's area: an fMRI study." *PloS one* 10, no. 11 (2015): e0141069.

Lerdahl, Fred, and Ray Jackendoff. A Generative Theory of Tonal Music. Cambridge, MA: MIT Press, 1983.

Lerdahl, Fred. Tonal Pitch Space. New York: Oxford University Press, 2001.

Maess, Burkhard, Stefan Koelsch, Thomas C. Gunter, and Angela D. Friederici. "Musical syntax is processed in Broca's area: an MEG study." *Nature neuroscience* 4, no. 5 (2001): 540-545.

Mozart, Wolfgang Amadeus. *Sonata No. 12 in F Major* K. 525B Bryn Mawr, PA. Theodore Presser Company, 1960.

Mukherji, Somangshu. "Generative Musical Grammar---a Minimalist Approach." (PhD Diss. Princeton University 2014). Print.
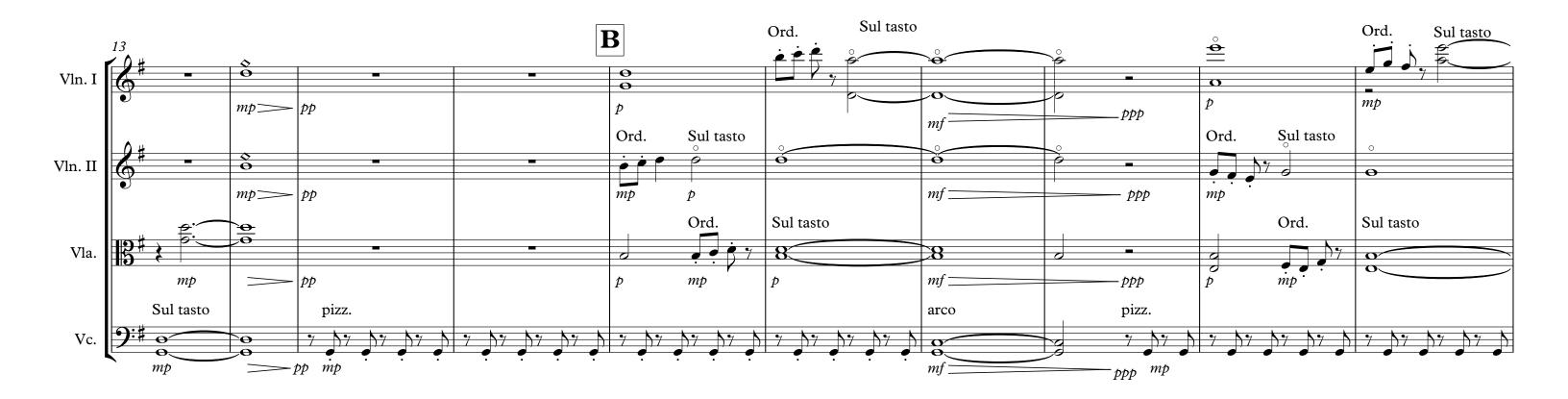
Marr, David. Vision: a computational investigation into the human representation and

processing of visual information. Cambridge, MA: MIT Press, 2010.

Narmour, Eugene. Beyond Schenkerism: The Need for Alternatives in Music Analysis. Chicago:

University of Chicago Press, 1977.

Narmour, Eugene. The Analysis and Cognition of Melodic Complexity: The Implication-

realization Model. Chicago: University of Chicago Press, 1992.

Pastille, William A. "Heinrich Schenker, Anti-Organicist." 19th-Century Music 8, no. 1 (July

1984): 29-36. Accessed December 16, 2016. doi:10.2307/746248.

Patel, Aniruddh D. Music, Language, and the Brain. New York: Oxford University Press, 2010.

Patel, Aniruddh D. "Using music to study the evolution of cognitive mechanisms relevant to

language." *Psychonomic Bulletin & Review* (2016): 1-4.

Patel, Aniruddh D., Edward Gibson, Jennifer Ratner, Mireille Besson, and Phillip J. Holcomb.
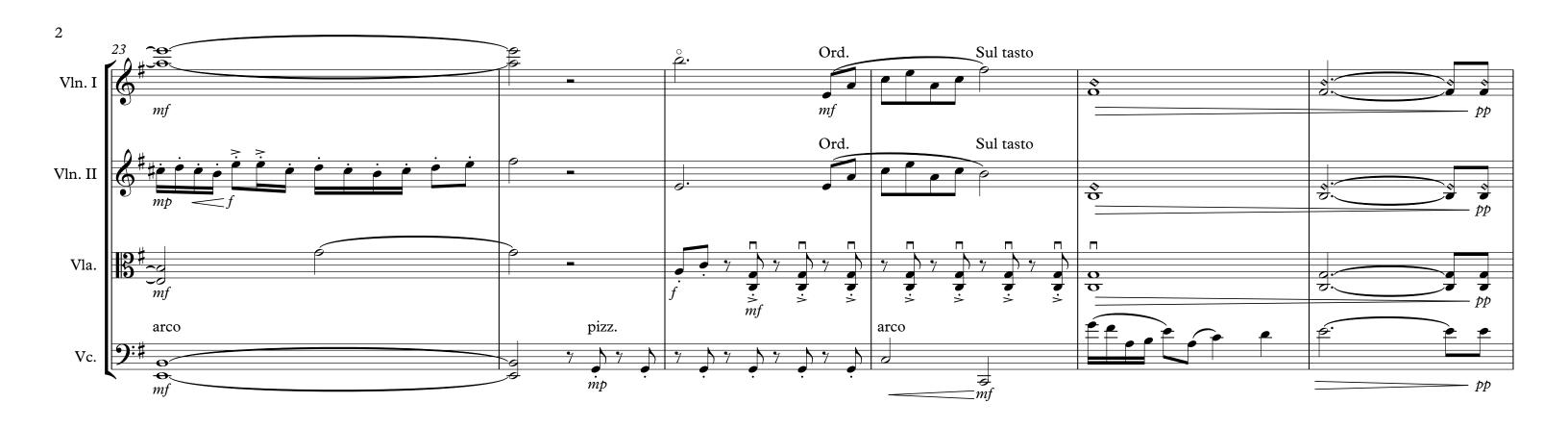
"Processing syntactic relations in language and music: An event-related potential study."
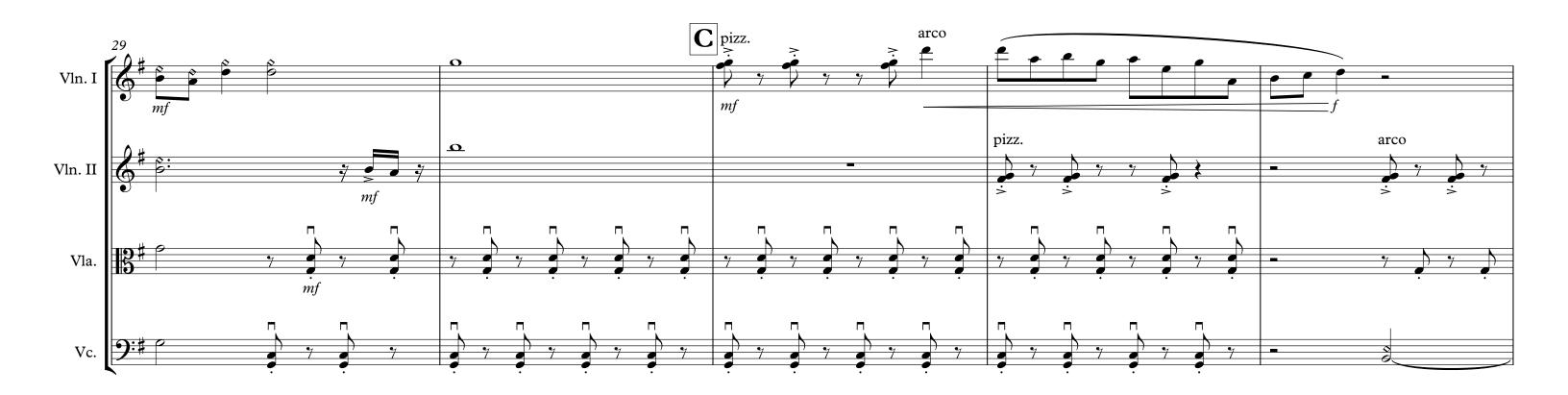
Journal of cognitive neuroscience 10, no. 6 (1998): 717-733.

Sammler, Daniela, Stefan Koelsch, and Angela D. Friederici. "Are left fronto-temporal brain

areas a prerequisite for normal music-syntactic processing?." *Cortex* 47, no. 6 (2011):

659-673.

Yang, Charles. "Ontogeny and phylogeny of language." *Proceedings of the National Academy of

Sciences* 110, no. 16 (2013): 6324-6327.
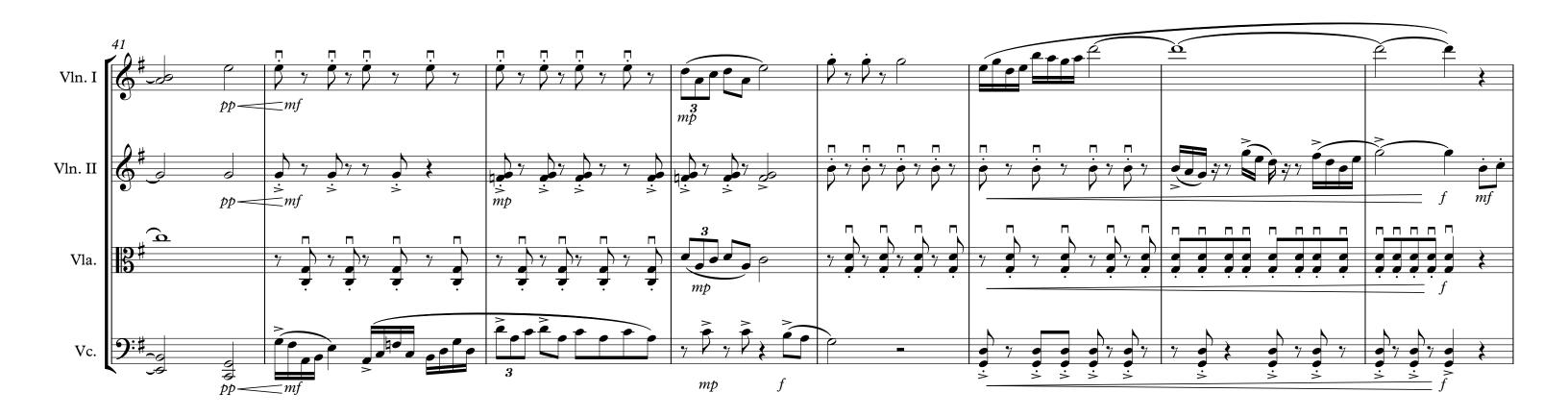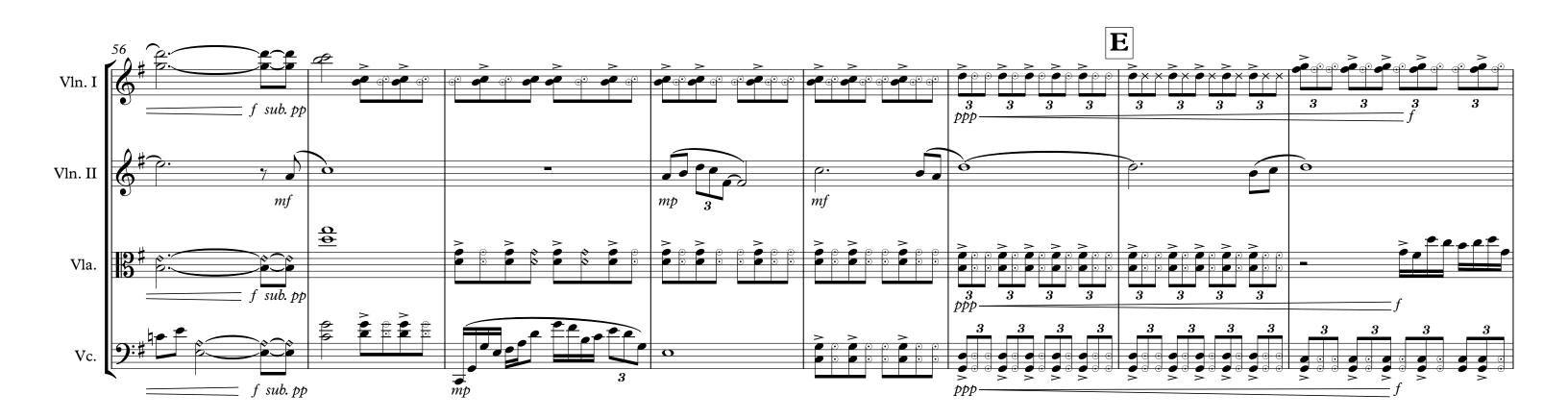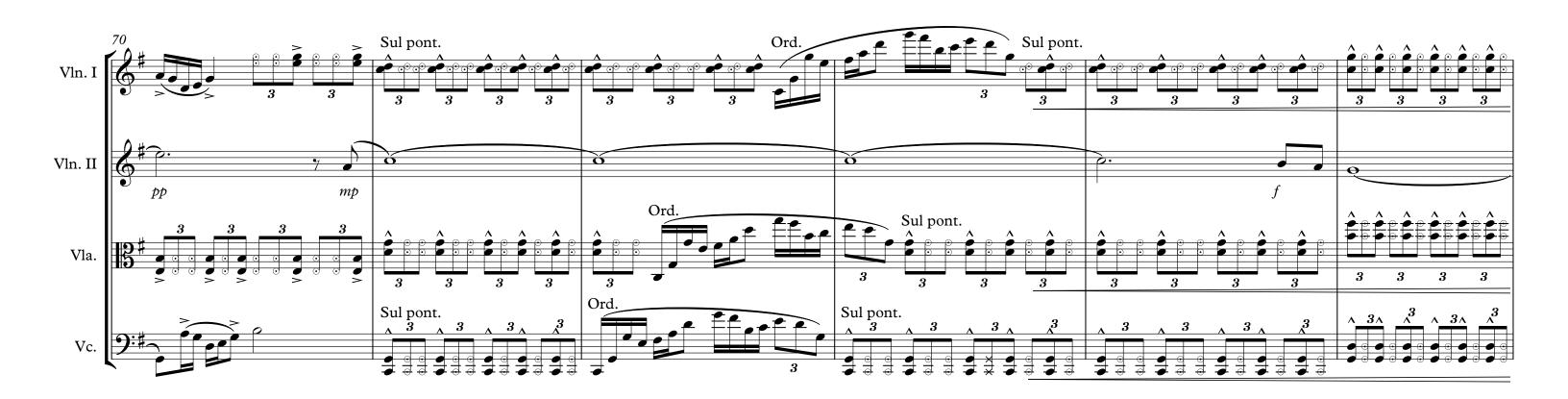
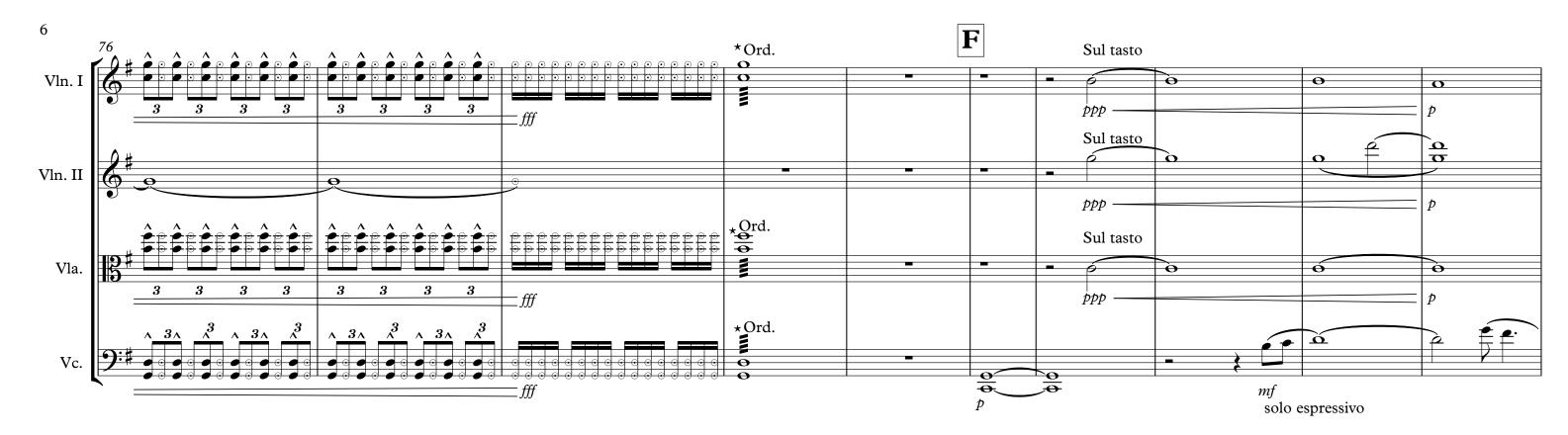# An Exploration of Musical Syntax:
Mvt. I: Hello

D

Vln. I — Col legno batuto⋆ — *mf* — *p*

Vln. II

Vla. — *mp* — Col legno batuto⋆ — *p*

Vc. — Col legno batuto⋆ — *mf* — *p*

⋆noteheads= pitch sounds
⊙ = slight pitch sound

E

Vln. I — *f sub. pp* — *ppp* — *f*

Vln. II — *mf* — *mp* — *mf*

Vla. — *f sub. pp* — *ppp* — *f*
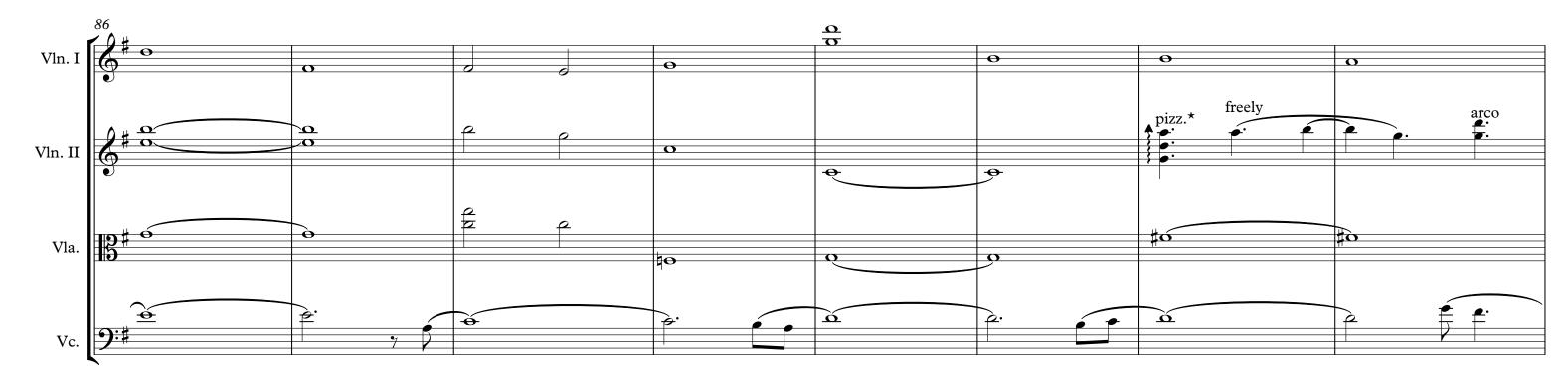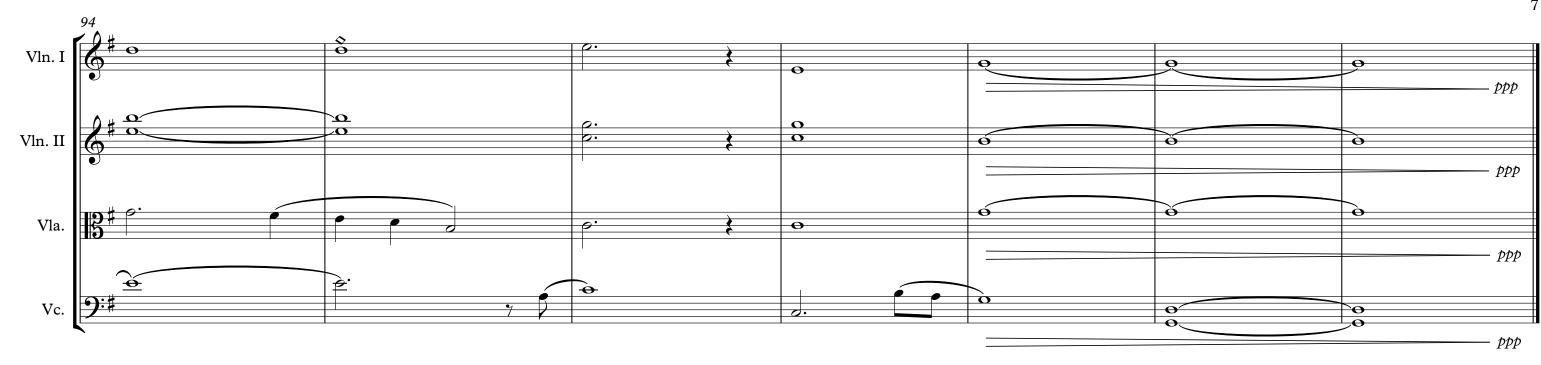
Vc. — *f sub. pp* — *mp* — *ppp* — *f*

F

* Note: Very fast tremolo, do not
worry about rhythm. Normal bowing,
stop suddenly.

* Note: Pizz. played with
Guitar-like strum

An Exploration of Musical Syntax
Mvt. 2: Forms

Allegro ♩ = 132

A

Violin I

Violin II

Viola

Violoncello

**Begin grace notes on the downbeat

B **Allegretto** ♩ = 106

like a walking bass

Pizz. chords played with
guitar-like strumming