Examining Contrasting Expressive Content within First and Second Musical Themes

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the Graduate School of The Ohio State University

By

Lindsay Alison Warrenburg, B.A.

Graduate Program in Music

The Ohio State University

2016

Thesis Committee:

David Huron, Advisor

Johanna Devaney, Co-advisor

Eugenia Costa-Giomi
Musical themes offer some of the most memorable parts of a musical work and have been a subject of interest in the study of music theory for centuries. In this thesis, contrasts between first and second musical themes are discussed, drawing from ideas in music theory and empirical musicology. The aim is to examine the construction of first and second musical themes. The thesis begins with a discussion of musical themes from the perspective of music theory. Next, a corpus study of over one thousand musical works was conducted, which chronicles differences between pairs of first and second themes. Features examined include the use of structural expressive features such as mode, rhythmic smoothness, interval size, articulation and dynamic markings, and durational pace. The results of the corpus study are consistent with the hypothesis that musical themes that appear first in a piece of music are more likely to be strong or energetic than are musical themes that appear second. In order to test whether listeners are sensitive to these structural differences, a perceptual study was conducted to examine whether musicians could discriminate between pairs of musical themes outside of the original musical context. The study was designed as a two-alternative forced-choice task and used both recorded and notated excerpts. The results are consistent with the hypothesis that listeners are able to differentiate between first and second themes using surface level features of the musical stimuli. The results of the task therefore provide some insight into
which musical and acoustic factors are perceived by musicians in both visual and auditory settings. The study concludes with a discussion about how the results might contribute to the conversation about musical themes from the perspective of music theory and musicology.
Acknowledgements

I would first like to express my gratitude to Drs. David Huron and Johanna Devaney. They have supported me throughout this process with patience, enthusiasm, and wise advice. I would also like to thank Dr. Eugenia Costa-Giomi, as a reader of this thesis. I am grateful for her valuable comments. I would also like to acknowledge Dr. Lawrence Bernstein of the Department of Music at the University of Pennsylvania, as he is the person that inspired me to study music as a career. His constant support over the years has made this research possible. Finally, I would like to thank my parents and my sister for their endless encouragement.
Vita

June 2009 .............................................. Academy of Allied Health and Science

2013 .............................................. B.A. Music, University of Pennsylvania

2014 .............................................. Research Technician, Monell Chemical Senses Center

2015 to present ................................ Graduate Teaching Associate, School of Music, The Ohio State University

Fields of Study

Major Field: Music

Music Theory

Music Cognition and Perception
# Table of Contents

Abstract .............................................................................................................................. ii  

Acknowledgments ........................................................................................................... iv  

Vita ........................................................................................................................................ v  

List of Tables ...................................................................................................................... vii  

List of Figures .................................................................................................................. ix  

Chapter 1: Introduction ....................................................................................................... 1  

Chapter 2: Considerations from Music Theory ................................................................. 5  

Chapter 3: Tests of Contrasting Expressive Content between First and Second Musical Themes ................................................................................................................ 21  

Chapter 4: Perception of First and Second Themes from Musical Scores and Sound Recordings ........................................................................................................................................ 57  

Chapter 5: Conclusions ..................................................................................................... 89  

References ........................................................................................................................ 95  

Appendix A: Evaluation of Barlow and Morgenstern’s Themes ......................................... 102  

Appendix B: Additional Materials from the Perceptual Study ........................................ 127  

Appendix C: Participant Scores and Quantitative Indices of Musical Works ................. 134
List of Tables

Table 1. Partial Reproduction of Table from Horn and Huron (2015) ..................24
Table 2. Most Frequent Key Relationships Between First and Second Thematic Pairs..............................................................................................................49
Table 3. Results of Statistical Analyses for Works Exhibiting One Introductory Theme, One First Theme, and One Second Theme ........................................54
Table 4. Criteria from the Corpus Study ................................................................62
Table 5. Categorization of Musical Stimuli Based on Their Articulation and Dynamic Levels ................................................................................................................65
Table 6. Arrangement of Musical Stimuli (n = 22) as a Cross of Musical Category and Condition .............................................................................................................69
Table 7. Summary Data for the Piano Works in the Expected, Opposite, and Two Conflicting Categories ........................................................................................................80
Table 8. Rank-Ordered Criteria Used to Determine Which Theme From a Pair Comes First or Second in the Music ............................................................................83
Table 9. Additional Criteria Listed by Participants as Features that Helped Them Perform the Discrimination Task .................................................................84
Table 10. A Comparison of Thematic Content in Musical Works .......................103
Table 11. A List of Musical Works Used in the Perceptual Study .........................127
Table 12. Descriptions of Four of the Piano Works in the Perceptual Tasks ......138
Table 13. Accuracy in Judgments About First and Second Themes in the Modified Perceptual Task ................................................................. 138

Table 14. Participant Scores and Quantitative Indices of Musical Works .......... 143
List of Figures

Figure 1. First Theme (top line) and Second Theme (bottom line) of Dvořák’s *Waltzes Op. 54, No. 3* .................................................................6

Figure 2. Rectangular Representation of the Surface of the Torus ...............40

Figure 3. Second Themes are Less Likely to Be In the Minor Mode Than Are First Themes .................................................................42

Figure 4. Second Themes Exhibit a Slower Pace Value Than Do First Themes...43

Figure 5. Second Themes Exhibit Quieter Dynamic Levels Than Do First Themes .........................................................................................44

Figure 6. Second Themes Are More Legato Than Are First Themes ..........45

Figure 7. First and Second Themes Do Not Differ in Terms of nPVI Values......46

Figure 8. Second Themes Utilize Slightly Smaller Intervals Than Do First Themes .........................................................................................50

Figure 9. Comparison of Modality (top graph) and Pace Value (bottom graph) In First and Second Themes Across Musical Periods .........................53

Figure 10. The Same Musical Score Shown in the “Without Dynamics and Articulation” Condition (top line) and in the “With Dynamics and Articulation” Condition (bottom line) ........................................................................68

Figure 11. Overall Judgments of First and Second Musical Themes ............75
Figure 12. Participants Perform Slightly Better in Situations When the Music Contains Articulation and Dynamic Information (“With D.A.”) Than When The Music Does Not Contain this Information (“Without D.A.”) ........................76
Figure 13. Participants Were Able To Distinguish First and Second Themes in The Score-Based and in the Recording-Based Task .................................77
Figure 14. Participant Performance Across the Four Categories of Thematic Stimuli .....................................................................................................79
Figure 15. First theme (top line) and Second Theme (bottom line) from Dvorak’s Waltzes, Op. 54 No. 3 ..................................................................................85
Figure 16. Bach’s Well Tempered Clavier Book 2, Prelude 3 .......................106
Figure 17. Handel’s Concerto Grosso in D Minor No. 10, Mvt 5 .....................107
Figure 18. Beethoven’s Piano Sonata Op. 2, No. 2, Mvt 1 ..............................108
Figure 19. Liszt’s Transcendental Etude No. 4 “Mazeppa” .............................109
Figure 20. Mozart’s Piano Concerto in Eb, K. 482, Mvt. 1 ..............................113
Figure 21. Score-Based Example With Accompaniment (No Dynamics/Articulation) ........................................................................................................131
Figure 22. Score-Based Example Without Accompaniment (With Dynamics/Articulation) ..............................................................................................132
Figure 23. Criteria Options for the Score-Based Task ....................................133
Figure 24. Beethoven’s Piano Sonata No. 22 in F Op. 54, Mvt 1 .................139
Figure 25. Brahms’ Rhapsody in E-Flat Op. 119, No. 4 ...............................140
Figure 26. Dvořák’s Waltzes Op. 54, No. 3 ....................................................141
Figure 27. Prokofiev’s Sonata No. 6 Op. 82, Mvt 1 .........................................142
Chapter 1: Introduction

Musical themes are short and distinctive sections of music, which listeners tend to find compelling and memorable. Within Western art music, musical themes are often considered to be fundamental elements of the composition. The role of musical themes within a piece of music is both multifaceted and variable, and carries with it aesthetic and structural implications. The deployment and development of musical themes within a single piece of music and across multiple musical works has been a topic of seminal importance in music scholarship. The investigation of musical themes can help determine the structure and evolution of a musical work. It can also help to fit the musical work within a wider historical and stylistic context.

Although musical works may contain any number of themes, it is common for them to employ two principal themes. In such works, the attributes of the two themes tend to contrast. These contrasts have been described in scholarly writings for centuries. In the eighteenth and nineteenth centuries, contrasting first and second themes were discussed by scholars such as Abbé Georg Joseph Vogler, Francesco Galeazzi, and Adolph Bernhard Marx. In the twentieth and twenty-first centuries, preeminent scholars like Charles Rosen, William Caplin, James Hepokoski, and Warren Darcy have examined differences in first and second themes extensively. Over this period of three hundred years, the methodologies used to study musical themes differed. Despite differences in
opinion and methodologies among theorists who utilized these different strategies, there exists a broad consensus regarding the expressive tendencies of first and second musical themes.

In order to examine how first and second musical themes differ, most scholars rely on standard musical analysis. There are many advantages to conducting close analysis; for example, when determining the placement of musical themes, it is possible to take the overall form and harmonic structure of a work into account. However, analyzing a large number of musical themes by hand takes many hours of work. Although it is possible to manually analyze a large body of musical works, these projects often take a number of years to complete. Other analytic musical observations arise from a more intensive examination of a small number of works. However, these works are sometimes selected without regard to the representativeness of the complete oeuvre. It could be that the music selection is motivated by pre-existing conceptions, whether or not the theorist is aware of this possibility. This raises the question of the generality of such an analytic claim. In summary, the use of traditional music theory methodologies tends to result in a trade-off: it is either difficult to provide a representative sample of musical works or it will take many years to complete a project.

The current study uses an empirical methodology in order to address the two concerns that arise from traditional music theory analysis. The thesis is a modest attempt to examine specific properties of musical themes in Western art music. Approaches from the areas of music theory, empirical musicology, and psychology will be integrated to investigate certain structural characteristics of musical themes, as well as how these characteristics are perceived by musicians. The hope is to contribute to the centuries-long
discussion regarding musical themes and to suggest productive avenues for further research.

Chapter 2 summarizes the argument that first and second themes contrast. In the late eighteenth and early nineteenth centuries, theorists explained these contrasts by using descriptive and metaphorical terms. Common descriptions of first themes included terms like *masculine*, *strong*, and *energetic*, whereas typical descriptions of second themes included terms such as *feminine*, *lyrical*, and *gentle*. Instead of relying on descriptive adjectives to explain thematic contrast, twentieth and twenty-first century scholars tend to examine the musical construction of themes. They describe contrasts between themes in terms of structural features like modality, rhythmic smoothness, articulation, and dynamic markings. Despite these different foci, there exists a broad agreement regarding the expressive tendencies of first and second musical themes. Discussions of thematic contrasts have entered into standard music theory textbooks and have been widely taught for over a century.

The study presented in Chapter 3 tests these classic theoretical claims about first and second themes. In order to assess these descriptive claims about thematic contrast, it is necessary to transform theoretical concepts into quantifiable data. In other words, it is necessary to operationalize what is meant by descriptive terms like *energetic* and *lyrical*. We will examine certain structural features of 1063 themes taken from Barlow and Morgenstern’s *Dictionary of Musical Themes* (1948). The experimental findings regarding the durational pace, dynamic levels, and articulations are consistent with traditional textbook descriptions that first themes are *energetic* and second themes are *lyrical* in nature, given our operationalizations.
Although the study of notated music does indeed reveal structural features that are consistent with the notion of thematic contrast, it does not address whether performers and listeners are sensitive to, or aware of, these contrasts. Chapter 4 presents a perceptual study that assesses whether musicians can perceive these differences. The study uses excerpts from musical scores and sound recordings, presented both with and without dynamics and articulation. The results of the perceptual study are consistent with the hypothesis that people can differentiate first and second musical themes, even when the themes are presented outside of the original musical context. In Chapter 5, the results of both studies are summarized and further observations and discussion are presented.
Chapter 2: Considerations from Music Theory

The aim of this chapter is to review various ideas and theories concerning first and second musical themes proposed by music scholars. In musicological and music theoretic writings over the past 300 years, contrasts between themes have been described in two ways: either through the use of descriptive adjectives and metaphors, such as strong themes and lyrical themes, or by comparing how each theme is constructed. Specifically, themes have been compared in terms of their modality, average interval size, rhythmic smoothness, articulation, and dynamic markings. To anticipate the conclusions of this chapter, we will see that theorists in the eighteenth and nineteenth centuries often described contrasting musical themes using descriptive terms; conversely, more modern theorists tend to refer to how the structure of themes can differ. Within recent undergraduate textbooks, both descriptive terms and structural considerations are used to describe contrasts between themes.

For centuries, first and second themes have been treated as contrasting musical ideas, with opposing emotional capacities and structural elements. First themes are commonly described as being energetic in nature, and comparatively stronger than the second theme.¹ Second themes are typically characterized as comparatively more gentle in nature, which “relieves the heated commotion and bolsters the ear with a pleasing

¹ This type of statement was seen in writings by Koch (Koch, 1793/1983), Galeazzi (Churgin, 1968), Kollmann (Kollmann, 1799), and Vogler (Newman, 1963).
contrast” (Vogler 1779, in Newman, 1963, p.34). An example that illustrates this contrast is Antonín Dvořák’s *Waltzes Op. 54, No. 3*, shown in Figure 1. The differences in nature between these two themes are demonstrated by their conflicting dynamic markings and articulations. It is evident that there is a clear dynamic contrast between the two themes; the first theme has forte accents, whereas the second theme is marked entirely pianissimo. The first theme also contains short rests and staccato markings, whereas the second theme contains a slur. Because of these structural differences, the first theme may evoke a more energetic character, whereas the second theme may evoke a more gentle character.

![Figure 1. First theme (top line) and second theme (bottom line) of Dvořák’s *Waltzes Op. 54, No. 3*](image)

The term ‘theme’ was used to describe music as early as the seventeenth century (Dunsby, 2002). However, it wasn’t until the eighteenth and nineteenth centuries that

---

2 Similar characterizations of second themes can also be seen in writings by Koch (Koch, 1793/1983), Galeazzi (Churgin, 1968), Kollmann (Kollmann, 1799).
themes became a popular topic in scholarly writing (Dunsby, 2002). Charles Rosen, in *The Classical Style*, claims that there were certain stylistic norms in music of the Classical era that led to the use of the term ‘theme’ by theorists in the eighteenth and nineteenth centuries. Some recent music scholars have stated that European composers during the nominal Classical era wrote music in a conventionalized style, leading composers to treat musical themes in similar ways (Caplin, 1998; Hepokoski & Darcy, 2006; Rosen, 1971). Therefore, eighteenth and early nineteenth century scholars were better able to discuss commonalities among themes than their predecessors.

Hepokoski and Darcy (2006) provide a conjecture for why Classical norms, like the use of contrasting first and second themes, may have arisen in the first place. They posit that the form of many Classical works directly reflects principles of the Enlightenment (Hepokoski & Darcy, 2006). For example, *clarity* was highly valued in Enlightenment thinking. Rosen, Hepokoski, and Darcy argue that clarity is reflected in Classical music through the use of clear divisions between large formal sections. Another important principle in Enlightenment thought is *balance*. Within Classical music, Rosen, Hepokoski, and Darcy argue that balance is represented through the symmetrical relation between large formal sections. Rosen (1971) contended that composers are able to achieve the qualities of clarity and balance in the music by composing each section in a distinctive manner.

Musical themes are thought to highlight these distinctive sections (Caplin, 1998; Hepokoski & Darcy, 2006; Rosen, 1971). Caplin posits that composers can use contrasting themes to help elucidate the general structure of the music. Rosen, Hepokoski, and Darcy postulate that the major formal divisions of the work can be more
clearly articulated through the use of contrasting themes. To support this argument, Hepokoski and Darcy point to notes found in some of Beethoven’s sketches. In some of his works, Beethoven refers to a secondary theme as “m.g.,” a possible abbreviation for *Mittel-Gedanke* – “middle thought” in German (Hepokoski & Darcy, 2006, p.119, footnote 4). Hepokoski and Darcy claim that this is evidence that composers may have thought about thematic contrast when writing musical works.

If Classical-era composers wrote music in ways that highlighted contrasts between sections, it raises the question if contemporary listeners were able to hear these contrasts when listening to performance. A few scholars believe that Classical-era musicians attended to contrasts during musical listening, and that these contrasts may have been used to help a listener understand the musical work (Hepokoski & Darcy, 2006; Rosen, 1971). Rosen specifically claims that contemporary listeners were able to follow the musical structure during concerts by listening for contrasting themes (Rosen, 1971). He believes that when audience members direct their attention to themes, they are better able to keep track of where they are in the music. Some preeminent scholars, such as Gjerdingen, Hepokoski, Darcy, and Rosen, believe that listeners intuitively compared musical works to an expected “norm” (Gjerdingen, 2007; Hepokoski & Darcy, 2006; Rosen, 1971). According to these historical conjectures, eighteenth and nineteenth century audience members came to concerts with well-formed expectations about how the music should be structured.

---

3 Rosen attributes this ability to the rise of public concerts in the 1780s (Rosen, 1971), arguing that it was due to this rise in public concerts that amateur musicians heard Classical music more frequently. It was through repeated listening to Classical music, Rosen believes, that audience members came to understand Classical stylistic norms.
Ideas about how themes contrast within musical compositions were present in writings by music theorists during the Classical era. The way theorists thought about musical composition during the eighteenth century has been described by Nicholas Cook (2002). He mentions that rather than using ideas of music theory to aid in the analysis of individual compositions, theorists approached music theory in a more observational way. Theorists are thought to have searched for commonalities among many musical works, for example the use of contrasting themes. If many musical works contained similar types of contrast between themes, then theorists would consider these contrasts to be important elements of music theory.

Hepokoski and Darcy (2006) believe that Classical-era theorists were likely to focus on the binary opposition between first and second themes, instead of on a more continuous musical experience. Ideas about differing thematic content in first and second themes were seen in writings by Vogler, Riepel, Galeazzi, and Kollmann (Hepokoski & Darcy, 2006). A particularly telling description of the theme types was given by Koch (1793), who described the primary theme as “a somewhat noisy/boisterous theme” and the secondary theme as “a cantabile theme” (Hepokoski & Darcy, 2006, p.118). As three examples of how Classical-era themes were discussed during this time, main ideas seen in writings by Jérôme-Joseph de Momigny, Anton Reicha, and A.B. Marx will be compared. Although these scholars primarily wrote about themes during the early and

---

4 Koch placed emphasis on the interdependence of key and thematic content, where he thought that the switch to a secondary key was necessary for second themes.  
5 For more details, the readers are encouraged to examine Scott Burnham’s excellent chapter on this history in his chapter “Form” in the *Cambridge History of Western Music Theory* (2002).
mid-part of the nineteenth century, their work mainly refers to music written during the nominal Classical era.

Jérôme-Joseph de Momigny, a French theorist writing at the outset of the nineteenth century, was preoccupied with thematic character and the role that musical themes played in the context of a work. As Burnham has noted, Momigny’s prioritization of melodic themes above their harmonic context and phrase structure was seen in writings by many theorists of this period, with musical personality becoming a central feature of analysis. Anton Reicha, a French theorist writing in the 1810s-1820s, was also preoccupied with differences in musical themes. He elevated the importance of the theme by discussing how listeners hear these *idées musicales* throughout an entire musical work: musical “theme[s] or motive[s] speak to our sentiment, flatter our ear, can be retained easily, and inspire the wish to be heard again” (Burnham, 2002, p.885).

A. B. Marx was a theorist of seminal importance in discussions about thematic function and character. Famously, Marx characterized sonata form as a type of organism, where “its subsections are not individual organisms but rather begin to function as interdependent and indispensable organs of a larger organism (‘the whole form’)” (Burnham, 2002, p.887). He related this thought to musical themes, indicating that that the two major theme groups in a sonata exposition experience a strong relationship: the first theme is in need of the second theme, and the second theme is dependent upon the first theme. In *Die Lehre von der musikalischen Komposition*, Marx

---

6 Reicha was especially interested in thematic use and development throughout a musical work, apart from their initial iterations in the piece.

7 Marx coined the standard German terminology for themes in sonata form. He labeled the first theme *Hauptsatz* (primary theme) and the second theme *Seitensatz* (secondary theme) (Burnham, 2002; Hepokoski and Darcy, 2006).
was the first to describe primary and secondary themes as gendered, with primary themes considered more “masculine” and secondary themes considered more “feminine” (Hepokoski & Darcy, 2006; Marx, 1845).

To conclude the discussion of how music theorists viewed contrasting musical themes during the eighteenth and early-nineteenth centuries, we can refer to a statement made by Galeazzi regarding second themes. He notes that second themes tend to be introduced “for the sake of greater beauty [and are typically] gentle, expressive, and tender in almost all kinds of composition” (Hepokoski & Darcy, 2006, p.132). Notice that the statement relies on metaphoric descriptions of themes. Terms like ‘gentle’ and ‘tender’ seem to relate qualities of themes to characteristics of people. These descriptive terms also appear in writings by other theorists of the time (Newman, 1963). Indeed, Galeazzi, Momigny, Reicha, and Marx all discussed thematic contrasts in Classical music with regard to the sentiments felt to be expressed by the respective theme.

Hepokoski and Darcy (2006) indicate that the dichotomy between energetic first themes and lyrical second themes may have been even stronger in Romantic music than in Classical music. Charles Rosen claims that during the Romantic period “second themes are generally more relaxed in every way than the first, with the [first theme] so unstable emotionally that a decrease in tension was inevitable” (Rosen, 1971, p.70). He specifically mentions the works of Schumann and Chopin as paradigmatic examples of music with contrasting themes. Within sonata form, Hepokoski and Darcy point to works by Schumann and Brahms that contain a de-energizing transition after a blocked or suppressed medial caesura. They speculate that devices like de-energizing transitions may have been used as a way to prolong the gentle approach to the second theme
(Hepokoski & Darcy, 2006). Therefore, the use of a *de-energizing transition* is thought to convey the contrast between the first and second themes in the sonata exposition.

In their book *Elements of Sonata Theory*, Hepokoski and Darcy indicate that widespread claims about themes during the Romantic era were likely influenced by significant nineteenth-century compositions. Recall that Marx made a statement in 1845 regarding “masculine” first themes and “feminine” second themes. Hepokoski and Darcy argue that Marx’s original statement about gendered themes was intended as a simple poetic description rather than as a serious claim. Nevertheless, some of the public latched onto the gendered terminology and came to think about first and second themes as gendered. Part of the reason why these gendered terms became popular was because of the prevalence of works such as Weber’s Overture to *Der Freischütz* and Beethoven’s *Coriolan* Overture, where the second themes are thought to refer to female characters. As a result of these seminal compositions, nineteenth century theorists may have expected to hear gendered first and second themes in works written by Classical composers. Indeed, Hepokoski and Darcy claim that theorists indiscriminately applied ideas about gendered themes from the nineteenth century to works written earlier. They state that although masculine primary themes (in their words, “stormy, threatened, troubled”) and feminine secondary themes (in their words, “eroticized or idealized”) do exist within the middle and later parts of the nineteenth century, it is less appropriate to apply these terms to any works before 1825 (Hepokoski & Darcy, 2006, p.147). Therefore, it seems that although there are clear contrasts between *energetic* first and *lyrical* second themes in both the Classical and Romantic periods, applying loaded descriptive terms such as “masculine” and “feminine” to Classical themes may not be appropriate.
Hepokoski and Darcy criticized the widespread use of gendered terms to describe musical themes because they believe that gendered terms only apply to themes in a few musical works. Namely, they critiqued scholars who used preconceived notions about gendered themes in order to interpret other musical works. Other preeminent scholars, such as Charles Rosen (1971) and Nicholas Cook (2002), agree that nineteenth century music theorists may have approached analysis with a predetermined view of what should happen in the music. Because of this approach to analysis, Rosen, Hepokoski, and Darcy have been critical of theoretical writings from the nineteenth century.

Scholars in the late twentieth and early twenty-first centuries have used alternative ways to characterize potential contrasts between themes. Instead of using descriptive terms such as strong or lyrical to describe themes, more recent scholars are primarily interested in examining the structural contrast between musical themes. This includes descriptions of how themes differ in terms of their construction. Features like modality, rhythmic smoothness, pace value, articulation, and dynamics become central ways to compare first and second themes.  

Additionally, an important distinction is made in recent scholarship between thematic construction and thematic function. The construction of musical themes refers to structural aspects, such as their melodic and rhythmic content. On the other hand, the

---

8 Structural features of first and second themes have been characterized by Caplin (1998) and Hepokoski and Darcy (2006). In a single musical work, first and second themes are not expected to differ with regard to all of these characteristics; they may differ with regard to only one or two of these categories. Nevertheless, the following structural features are thought to differ between first and second themes: Modality (Caplin, 1998, p.119, 197, 213; Hepokoski & Darcy, 2006, p. 119), Rhythmic Smoothness (Caplin, 1998, p.197, 201, 213, 229; Hepokoski & Darcy, 2006, p. 66, 132), Pace (Caplin, 1998, p.100, 113), Articulation (Hepokoski & Darcy, 2006, p.133), Dynamics (Caplin, 1998, p.123, 197; Hepokoski & Darcy, 2006, p. 66, 132).
function of a theme depends on its role in the music. In the case of sonata form, many theorists argue that the function of the secondary theme is to establish the subsidiary key (Caplin, 1998; Hepokoski & Darcy, 2006; Rosen, 1971). The construction of the secondary theme may consist of features like softer dynamic markings and slurs (Hepokoski & Darcy, 2006, p.133). Music scholars have placed different emphasis on the importance of thematic construction and thematic function. Schoenberg emphasized structural features of themes, such as their rhythmic and intervallic features, the metrical context, and their harmonic progressions (Dunsby, 2002), using these features to differentiate first and second themes (Caplin, 1998; p.97). On the other hand, Caplin primarily discusses themes in terms of their formal functions – how the music is able to express its location in musical time. Other music scholars have given more equal weight to thematic construction and function, as seen in the work of James Hepokoski and Warren Darcy.

Twentieth and twenty-first century theorists have described contrasts between first and second themes in terms of their musical construction. From his own musical analyses, Rosen (1971) concluded that before 1750, themes tended to differ in terms of their orchestration or voicing. He deemed these features “external musical features,” because they do not refer to the melody itself (Rosen, 1971). However, Rosen noticed that Classical era music contained different types of contrast between themes. He noted that Classical themes tended to contrast in terms of their rhythm, mode, speed, texture, and dynamics. He called these features “internal musical features” (Rosen, 1971). Rosen

---

9 These structural features are used to determine whether a theme is loosely organized or tight-knit in organization. Schoenberg believed that first and second themes differed in their overall organization. A discussion of this occurs later in the chapter, with regard to Caplin.
came to the conclusion that internal contrasts between themes are “not only frequent, but essential to the [Classical] style” (Rosen, 1971, p.82).

Hepokoski and Darcy (2006) also speak about structural features in what they call the rhetorical goal of sonata expositions. They analyzed a large number of musical works in sonata form in order to “perceive the musical result sculpturally, in terms of how it realizes melodic pattern, meter, tempo, articulation, dynamics, timbre, density, drive-to-cadence, and the like” (Hepokoski & Darcy, 2006, p.614). To chronicle how often composers used these structural features, they offer a succession of default-levels. First-level defaults are the most common occurrences of an event and are thought to be the “norm” in Classical-era works. Hepokoski and Darcy explain thematic contrasts by listing various default-levels for first and second themes.

Within Hepokoski and Darcy’s (2006) conformant model of sonata form, the function of first themes is to initiate the rhetorical rotation and to express the emotional state or personality of the sonata as a whole. The first-level default for first themes is a “strong-launch,” where the theme tends to be aggressive or emphatic in nature. Structurally, first themes are typically forte, and tend to contain dotted rhythms, fanfare-like gestures, and octave leaps (Hepokoski & Darcy, 2006, p.66). The “strong-launch” option is therefore reminiscent of the way earlier theorists had thought about strong or energetic first themes. However, Hepokoski and Darcy also allow for a “weak-launch” option, where first themes can start with a piano dynamic and can be lyrical in nature (Hepokoski & Darcy, 2006, p.66).  

---

10 They note that this “weak-launch” typically occurs after a slow introduction, particularly within works by Haydn.
according to the performing instruments, with idiomatic defaults for piano music, chamber music, and orchestral music (Hepokoski & Darcy, 2006, p.68).

According to Hepokoski and Darcy’s theory, the role of the secondary theme is to act as a “relaunch” into the work (Hepokoski & Darcy, 2006). Hepokoski and Darcy remark that second themes are typically in a new key, quiet in dynamic, and often lyrical (Hepokoski & Darcy, 2006, p.132). Although they state that the first-level default for second themes is lyrical and beautiful, they allow second themes to belong to one of seven possible heuristic categories. They claim that the type of second theme used in a piece of music often depends on whether the first theme begins as a strong- or weak-launch. Therefore, in Sonata Theory, first and second themes are not required to contrast in the sense of energetic and lyrical themes. However, Hepokoski and Darcy still emphasize structural contrast between first and second themes, explaining differences in terms of mode, rhythmic smoothness, dynamics, and articulation.

Comparing Hepokoski and Darcy to Caplin is an interesting task. Caplin agrees that first and second themes can differ in terms of their construction, although his primary concern is with the function of the themes. Throughout his book Classical Form, Caplin mentions that first and second themes can differ in terms of their modality, rhythmic

11 The second-level default of secondary themes is a “bustling, staccato, energetically gallant, or jauntily self-confident S,” which consists of nervous quirks, light, strutting steps, and are typically piano but have forte outbursts (Hepokoski & Darcy, 2006, p.132). There are also P-based S themes, as seen in Haydn, as well as S themes that are “Contrasting Derivation from P,” “Forte S,” “S as Virtuosic Figuration,” and “Fugal/Imitative S” (Hepokoski & Darcy, 2006, pp.133-139).

smoothness, pace value, and dynamics.\footnote{\textit{Modality} (Caplin, 1998, p.119, 197, 213), \textit{Rhythmic Smoothness} (Caplin, 1998, p.197, 201, 213, 229), \textit{Pace Value} (Caplin, 1998, p.100, 113), \textit{Dynamics} (Caplin, 1998, p.123, 197).} Caplin mentions that these structural features are used to support form-functional articulations, which are primarily created through harmonic, phrase-structural, and cadential considerations (Caplin, 1998). For this reason, rather than mentioning how each specific structural feature – like mode or dynamics – differs between first and second themes, Caplin discusses contrast between first and second themes in terms of their overall phrase-structure and temporal expression.

However, he mentions that “the start of a new formal unit [like a theme] often brings new melodic-motivic ideas” (Caplin, Hepokoski, & Webster, 2010, p. 38-39). Each of the structural features mentioned by Hepokoski and Darcy can therefore play a role in determining the character of themes; however, in Caplin’s view, it is the combination of these structural features and their harmonies and phrase-organization that gives rise to a typical first or second theme.\footnote{Themes are organized so that small sections of the theme carry distinct functions. Within a single theme, there are four separate functions: that of presentation, continuation, cadential material, and closing. The way these small functions are used within a single theme depends on the higher-order function of the entire theme itself. For example, a primary theme has a different method of establishing presentation and continuation functions than does a secondary theme. A closing theme will give different weights to cadential and closing material than will primary themes, transitions, and secondary themes.}

First themes are described by Caplin as having a “tight-knit” construction. They contain prominent sections of tonic prolongation, as well as an overall feeling of stability. He divides second themes into two types: subordinate themes (for sonata-form works) and interior themes (for works in large ternary form). Both of these second theme-types are considered to be more loosely constructed than their paired first theme, and thus tend
Caplin refers to differences between first and second themes, where he mentions that “just as the subordinate key contrasts with the home key, so too does the subordinate theme contrast with the main theme” (Caplin, 1998, p.97). He describes interior themes as being opposite in mode to their paired first themes. In addition, he mentions that interior themes often contrast with the main themes in terms of their melodic-motivic material, texture, and rhythmic accompaniment (Caplin, 1998). Within the context of a minuet-and-trio, he mentions that the trio theme (the second theme) often contrasts with the minuet theme (the first theme). By his account, the trio theme is thought to “bring a quality of simplification and relaxation” to the work after the minuet theme (Caplin, 1998, p.229). In general, Caplin extends this discussion of contrast between first and second themes to music written in many forms of the Classical style.

In summary, Caplin, Hepokoski, and Darcy tend to examine thematic contrasts in terms of their construction. Commonly examined structural features include mode, dynamics, changes in surface rhythm, and changes in articulation. Although the focus of these recent theorists differs from that of eighteenth and nineteenth century theorists, there are similarities between the two groups. In fact, it seems like the structural features

\[\text{15} \text{ Caplin examines both the thematic content of themes – their mode, rhythmic smoothness, durational pace, and dynamics – and features that determine a theme’s formal function – such as harmony, cadence-type, and key. He believes that the way a first theme is organized is fundamentally different than the way a second theme is organized. The different organizations allow the two themes to serve distinct functions within the music. First themes, which are tight-knit, are tonally stable, end with an authentic cadence, have symmetrical phrase groupings, and a general unity of melodic-motivic material. On the other hand, second themes often have a loose organization and tend to be tonally unstable (modulate or in the subordinate key), have cadential evasion or omission, diversity of melodic-motivic material, chromatic alterations, and asymmetrical phrase groupings.}\]

\[\text{16} \text{ The reader is encouraged to turn to Appendix A, in which first and second themes from a Mozart piano concerto are compared from the viewpoint of Hepokoski and Darcy, Caplin, and Barlow and Morgenstern.}\]
mentioned by recent scholars are related to the descriptive characteristics written about by earlier music theorists. First themes are described metaphorically as *strong* or *energetic* by earlier theorists, and are described structurally as having features like louder dynamics and more staccato markings by Caplin, Hepokoski, and Darcy. Likewise, second themes are characterized by earlier theorists as being *lyrical* or *gentle*, and are characterized by recent theorists as being quieter and having smoother articulations. There is a clear connection between these descriptions of thematic contrast.

In today’s music theory textbooks, thematic contrast is described in both the metaphoric sense of the early theorists, as well as in the structural sense of later theorists. The discussion of musical themes and musical form was examined in three of the most popular introductory textbooks in music theory education: *Harmonic Practice in Tonal Music* (Gauldin, 2004), *Tonal Harmony with an Introduction to Twentieth-Century Music* (Kostka & Payne, 2004), and *The Complete Musician: An Integrated Approach to Tonal Theory, Analysis, and Listening* (Laitz, 2012). Within all three of these textbooks are descriptions about how first and second themes tend to contrast. Thematic contrasts tend to be introduced in passages describing binary or ternary form, where A and B sections are taught to have distinct and contrasting melodic profiles (Gauldin, 2004; Laitz, 2012). In these forms, the B section “can provide contrast with the A sections by using different melodic material, texture, tonality, or some combination of these” (Kostka & Payne, 2004, p.338). In discussions of sonata form, the second theme is explained as being “generally more regular in its phrase groupings, as well as quieter and more lyrical than the opening theme, often employing a smaller instrumental force in symphonies” (Gauldin, 2004, pp.563-564). Metaphors are similarly applied to these first and second
themes, where “primary themes frequently sound more vital, grand, or ceremonial than the secondary themes…. [which] frequently have a lyrical or gentle character” (Kostka & Payne, 2004, p.345).

Despite different approaches among theorists, there exists a broad consensus regarding the expressive tendencies of first and second musical themes. These descriptive characterizations are usually related to the expressive musical content of the themes, where first themes are labeled as strong, masculine, and energetic, and the contrasting second themes tend to be described as lyrical, feminine, or gentle. Having reviewed various ideas concerning first and second themes, the question arises of how we might evaluate their representativeness in a large body of works. Accordingly, for the remainder of this thesis, I will endeavor to assess one of the most common and venerable claims regarding thematic organization – namely, that first and second themes contrast.

Theoretical claims can be evaluated from at least two perspectives. First, are the purported features evident in the notated themes themselves? Second, are the purported differences salient to non-theorists, either when viewing the notation or when hearing the music? In this thesis, both approaches will be used. Chapter 3 describes a corpus study, which examines structural features of classical music themes over a period of 300 years. Chapter 4 describes two perceptual studies that examine non-theorists’ abilities to decipher contrasting themes when viewing musical notation and when hearing recordings. The results of these two tasks will be used to facilitate a discussion of how the implications from this research apply in the broader perspective of music theory.
Chapter 3: Tests of Contrasting Expressive Content between First and Second Musical Themes

Chapter 2 described that theorists have long noted a binary opposition between the first and second themes of Western art-music works. Early writings about contrasts between these themes often contained descriptive terms, such as strong and energetic first themes and gentle and lyrical second themes. On the other hand, modern scholars use thematic structure to describe contrasts between first and second themes. Generalizations about thematic contrast are also evident in current textbooks of music theory. The goal of this chapter is to test these classic theoretical claims about first and second themes. The study is motivated by two questions: (1) Can we observe the contrasts between first and second themes described in the extant literature? (2) Can we observe changes in the construction of first and second themes over different stylistic musical periods?

Hypotheses

As we have seen, the descriptive literature characterizes first themes using terms such as energetic and strong, whereas second themes are characterized using terms like lyrical, quiet, and gentle. In order to test these notions, we need to operationalize these descriptive terms in a way that permits some form of measurement.
Breaking down rich, descriptive terms into component parts is a subjective act. However, we can use the descriptive literature cited in Chapter 2 to help us perform this task. In order to describe first themes, eighteenth and nineteenth century theorists used descriptive terms like strong and energetic. Modern scholars, like Hepokoski, Darcy, and Caplin, have described structural features of first themes: loud dynamics, a fast pace, staccato or detached articulation, and disjunct rhythms. These structural elements were deemed appropriate operationalizing terms for the descriptions of first themes (strong, fiery, energetic). A similar trend can be seen in descriptions of second themes. In order to characterize second themes, Vogler, Galeazzi, and Koch have used terms like lyrical, quiet, and gentle. If second themes are written in a manner that is lyrical, quiet, and gentle, we might expect that they have quieter dynamics, a slower pace or tempo, legato articulation, and smoother rhythms. Once again, we can see that these are the structural features of second themes referred to by Hepokoski, Darcy, and Caplin.

One final theory might be made with regard to musical mode. Traditional music theory indicates that the second theme is more likely to be in the major mode than is the first theme.\textsuperscript{17} However, an alternative expectation arises when we consider the emotional mood generally associated with music in the major and minor modes. Major mode music is typically associated with happy or positive affect (Temperely & Tan, 2013). Given this

\textsuperscript{17} Within sonata form, it is typically assumed that the primary theme is in the tonic key (Caplin, 1998; Hepokoski & Darcy, 2006; Rosen, 1971). For works in the major mode, the secondary theme tends to be in the key of the dominant (Hepokoski & Darcy, 2006). According to Sonata Theory, for works in the minor mode, the secondary theme is most often in the key of the relative major, but can also be in the key of the minor dominant (Hepokoski & Darcy, 2006). Therefore, for works in the major mode, the primary and secondary themes would be expected to be in the major mode. For works in the minor mode, the primary theme would be expected to be in the minor mode, but the secondary theme would be expected to be in the major mode.
observation, we might expect that energetic first themes would be in the major mode. Music in the minor mode has been associated with quieter dynamics and a slower tempo (Horn & Costa-Giomi, 2011; Ladinig & Huron, 2010; Post & Huron, 2009; Turner & Huron, 2008). Therefore, we might expect second themes, which are thought to have quiet dynamics and a slower tempo, to be in the minor mode.

Work by Horn and Huron (2015) supports the notion of major-mode first themes and minor-mode second themes. The authors examined the use of mode between the years 1750 and 1900. In order to do this, they used cluster analysis to examine the relationships between modality, tempo, articulation, and dynamics. Table 1 shows a partial reproduction of their findings. From this table, we can see that across the years 1750-1900, there are consistently more works written in the major mode than in the minor mode. A trend emerges when we examine the relationship between tempo, articulation, and dynamics across works in the same mode (major or minor). Among major-mode works written between 1750-1849, there is a relatively equal occurrence of music that is “loud, fast, and staccato” and music that is “quiet, slow, and legato.”

Recall that we have categorized first themes as typically “loud, fast, and staccato” and second themes as typically “quiet, slow, and legato.” A possible interpretation of the data is therefore that the major mode might be used relatively equally between first themes (“loud, fast, and staccato”) and second themes (“quiet, slow, and legato”).

Among works in the minor mode, there is consistently a higher proportion of works that are “quiet, slow, and legato” than works that are “loud, fast, and staccato.”

---

18 Within the period 1850-1899, there is a higher proportion of major-mode works that are quiet, slow, and legato than major-mode works that are loud, fast, and staccato. This period might be considered to be an exception.
first themes are typically “loud, fast, and staccato” and second themes are typically “quiet, slow, and legato,” we might wonder if the minor mode would be used more frequently within second themes than among first themes. Recognizing that this differs from standard music theory, we might expect that second themes are more likely to be in the minor mode.

<table>
<thead>
<tr>
<th>Features</th>
<th>Characterization</th>
<th>1750-1799</th>
<th>1800-1849</th>
<th>1850-1899</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Major</em> + loud, fast, and staccato</td>
<td>Joyful</td>
<td>22%</td>
<td>16%</td>
<td>18%</td>
</tr>
<tr>
<td><em>Major</em> + quiet, slow, and legato</td>
<td>Tender/Lyrical</td>
<td>22%</td>
<td>20%</td>
<td>44%</td>
</tr>
<tr>
<td><em>Minor</em> + loud, fast, and staccato</td>
<td>Passionate</td>
<td>4%</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td><em>Minor</em> + quiet, slow, and legato</td>
<td>Sad/Relaxed</td>
<td>7%</td>
<td>14%</td>
<td>21%</td>
</tr>
</tbody>
</table>

Table 1. Partial reproduction of table from Horn & Huron (2015). Features of first themes are thought to be loud, fast, and staccato; features of second themes are thought to be quiet, slow, and legato. The characterization labels are taken from Horn & Huron (2015).

Accordingly, the descriptive terms about first and second themes were interpreted as having the following repercussions:

*First themes – Strong or energetic:* louder dynamic, relatively faster pace, staccato or detached articulation, major mode.

*Second themes – Lyrical, quiet, and gentle:* quieter dynamic, slower pace or tempo, legato articulation, minor mode.

These ideas can be expressed in the form of the following hypotheses. Compared with first themes, second themes are conjectured to be:
H1: more likely to be in the minor mode
H2: more likely to exhibit a slower tempo or pace
H3: more likely to have a quieter dynamic level
H4a: more likely to exhibit smoother (more isochronous) rhythms
H4b: more likely to be legato
H5a: more likely to be in a different key
H5b: more likely to be in a closely related key
H5c: more likely to be in the dominant key, or the relative major (if the first theme is minor), or the relative minor (if the first theme is major)

None of the extant characterizations of first and second themes makes a prediction regarding interval use. Nevertheless, we propose to test whether there is a difference in interval use between first and second themes—specifically regarding average interval size. *A priori*, it is not clear whether first or second themes might exhibit larger average melodic intervals. On the one hand, energetic themes might be expected to employ more melodic leaps. On the other hand, extant research has shown that large intervals tend to employ long durations,\(^\text{19}\) so the slower, presumably more lyrical second themes might exhibit larger intervals. Neither of these ideas sticks out as the more obvious conjecture. Consequently, we will simply test whether there is any difference in interval size between the two types of themes. Specifically, we will test:

H6: First and second themes exhibit different average interval sizes.

---
\(^{19}\) This is consistent with Fitts’ law. For more on Fitts’ law, see Huron (2001) and Temperley (2014, p.175-176).
Sample

In order to identify a large sample of musical themes, a convenience sample of encoded musical themes was used: the *Dictionary of Musical Themes* (1948) assembled by Harold Barlow and Sam Morgenstern. The *Dictionary* was conceived as a reference tool for identifying the source (composer and work) of Western art-music melodies or themes (Barlow & Morgenstern, 1948). The first edition of the *Dictionary* consists of some 9,788 themes from instrumental musical works spanning a period from the late sixteenth to the early twentieth centuries. It necessarily reflects the musical interests of Barlow and Morgenstern and so many musicologists would regard the *Dictionary* as biased. Indeed, there seems to be a bias toward nineteenth-century orchestral works, as well as a bias towards American composers in the twentieth century.

Critical for the purposes of this study, Barlow and Morgenstern labeled the order of themes within works, giving labels such as “first theme” and “second theme.” In addition, some themes were labeled as “introductory.” Barlow and Morgenstern also labeled certain programmatic themes, such as the “Peter” and “Duck” themes from Prokofiev’s *Peter and the Wolf*. These latter themes are concurrently labeled “A,” “B,” etc., reflecting their order of appearance. Throughout the thesis, the themes compared are labeled “first” and “second” in the *Dictionary*, which are determined by their order of presentation within the musical work. The developing treatment of these themes throughout the musical work is not considered. For example, occurrences of the same musical themes in the development or recapitulation of sonata form works are not investigated.

For the purposes of this study, sampled works were limited to those that contained
only two themes. In order to avoid confusion, we excluded any works that also contained a theme labeled “Introduction.” A total of 1063 works met this sampling criterion and were included in the main analysis. The sampled music included a wide variety of composers, stylistic musical periods, and musical genres (sonatas, symphonies, ballets, solo instrumental works, chamber works, etc.). However, no vocal music is included in the Barlow and Morgenstern Dictionary and so was not considered in this thesis.

A common thread among the writings of twentieth and twenty-first century music scholars is the sentiment that musical themes should be labeled through the process of close analysis rather than through the use of conventional labels (Hepokoski & Darcy, 2006; Caplin, 1998, Ratner, 1980, Rosen, 1971). Said differently, we should let the musical construction determine the true identity of themes within a musical work. An appropriate question, therefore, is how the Dictionary of Musical Themes is organized: how do the themes labeled “first” and “second” correspond with traditional musical analysis? This author has not found any documentation of how Barlow and Morgenstern determined the themes in their Dictionary. Because of this mysterious methodology, it

---

20 The Barlow and Morgenstern Dictionary has mainly been used in scholarly work as a “large and representative” convenience sample of Western art music (Simonton, 1989). In particular, it has been used in the fields of psychology and music information retrieval to address diverse questions such as how people recognize and remember melodies (Gardiner et al. 1996; Java et al. 1995), how auditory imagery arises from musical notation (Brodsky et al. 2003), how to create effective methods of music retrieval and highlight its difficulties (Byrd & Crawford, 2002; Doraisamy & Rüger, 2003; Downie, 2003), how to compare prosody and rhythm in language and music (Patel & Daniele, 2003; Patel et al. 2006; Huron & Ollen, 2003), and to examine the use of intervals in ascending and descending melodies (Vos et al. 1989). Dean Simonton conducted a series of studies using the Barlow and Morgenstern Dictionary as a way to measure the musical output of composers, where the number of “famous” themes – themes accounted for in the Dictionary – a composer wrote in certain periods was a central feature of interest (Simonton 1977, 1980, 1989, 1991, 2000). In these studies, the Dictionary was mainly used as a standard representation of Western art-music works.Although general remarks
is possible that musical “themes” – which carry a great functional weight within musical works – are simply “tunes” or “melodies” that may stick out to Barlow and Morgenstern as memorable and distinctive. Additionally, themes that are presented first and second in a musical work may have different formal designations through the lens of traditional music theory. For example, “first” and “second” musical themes may correspond in various ways to functions in sonata-allegro form: “first” and “second” themes could be indicators of a “primary” and “secondary” theme (the normative conclusion in this thesis), but may also correspond to a “primary theme 1” (first theme) and “primary theme 2” (second theme) or to a “primary theme” (first theme) and “transition” (second theme). For works not in sonata form, the “first” theme may correspond to a “minuet” theme and the “second” theme may be akin to a “trio” theme.

To allay concerns regarding the use of the Barlow and Morgenstern Dictionary, an analysis of a few musical works was conducted. This allows the musical context to be taken into account when determining the placement of musical themes, as well as the overall musical form and harmonic structure of the work. The reader is encouraged to turn to Appendix A, where the results of the close analyses are discussed. Overall, the themes from the theory-driven analysis correspond well with the themes from Barlow and Morgenstern’s Dictionary.

were made in many of these papers about the limitations of using a convenience sample like the Dictionary, the nature of these studies did not rely on specific correspondence between the themes in the Dictionary with the themes from an analysis of music theory. Therefore, to this author’s knowledge, there has been little or no effort to externally verify the themes in Barlow and Morgenstern’s Dictionary of Musical Themes.
Method

All of the materials were encoded in the Humdrum format and all processing was carried out using the Humdrum Toolkit (Huron, 1993). For each first and second theme, our hypotheses require operationalized measures for modality (major or minor), speed (tempo or pace), articulation (staccato to legato), rhythmic smoothness, dynamic level, key relationship, and interval size. The operationalizations of these measures are described below.

Mode

In the case of mode, Barlow and Morgenstern explicitly interpreted all of the themes in the Dictionary as being in the major or minor mode. No effort was made to verify their modal designations.

Pace

Hypothesis 2 requires some characterization of the tempo, speed, or pace of the first and second themes. The Dictionary includes no indications of tempo. Even if a work or movement includes an explicit tempo indication (either a metronome marking or an Italian tempo term) the likelihood is that tempo indication would apply to both the first and second themes. Although performers may tend to perform one of the themes faster than the other theme, no effort was made to examine the pertinent performance practice.

Instead, we might expect that the “pace” of a musical passage is likely to be reflected in the relative durations of the notes in the two different themes. That is, a composer might create a “slower” theme by simply notating the theme using notes of
longer duration. Accordingly, we propose to characterize the “pace” (as distinct from the tempo) of themes by attending to the average note lengths. Faster-paced themes are likely to exhibit more short duration notes (such as eighth and sixteenth notes), whereas slower-paced music often exhibits more long duration notes (such as half and whole notes).

There are a number of ways to characterize musical speed or pace. In a mixed sequence of long and short notes, short notes appear to play a greater perceptual role in defining speed or pace than long notes. Long notes often simply indicate moments of repose at the ends of figures or phrases. Consequently, a simple average of elapsed duration seems to exaggerate the influence of long notes. For example, the average of a whole note and a quarter note is significantly longer than a half note. A simple operationalization that emphasizes shorter notes can be achieved by averaging reciprocals of the conventional duration designations. If wholes, halves, quarters, eighths, etc. are represented by the numbers 1, 2, 4, 8, etc., then we can simply average these reciprocal values across all notes in the theme. In this operationalization, rests are excluded. By way of illustration, consider a 7-note theme containing four quarter-notes, two half-notes, and one whole-note. We might determine a “pace value” as follows: \( 4 \times 4 + 2 \times 2 + 1 = 21; \frac{21}{7} = 3.0 \). Similarly, a 10-note theme containing six sixteenths, three eighths, and one quarter would have a pace value of \( 6 \times 16 + 3 \times 8 + 1 \times 4 = 124; \frac{124}{10} = 12.4 \). In general, themes with shorter note durations will produce higher pace values.

The concept of pace value can be illustrated by comparing three similar rhythms, each in 3/4 meter. The first rhythm is ‘half quarter half quarter.’ The pace value for this rhythm would be \( 2 \times 2 + 2 \times 4 = 12; \frac{12}{4} = 3 \). The second rhythm adds notes on the

---

21 For more on this topic, see Broze & Huron (2013).
second beat resulting in six quarter notes. The pace value for this rhythm would be \((6 \times 4) = 24\); \(24 \div 6 = 4\). The third rhythm modifies the first rhythm so each quarter note is replaced by two eighth notes: ‘half eighth eighth half eighth eighth.’ The pace value for this rhythm would be \((2 \times 2) + (4 \times 8) = 36\); \(36 \div 6 = 6\). Notice that the second and third rhythms contain the same number of notes and so would exhibit the same average duration. However, the pace values differ, reflecting the greater activity due to the presence of the shorter eighth notes. Both the number of notes, as well as the duration of each note, play into considerations of durational pace. The change in durational pace from a first theme to its paired second theme was determined by calculating the difference between the durational pace for the first and second theme.

**Measures of Musical Dynamics**

Unfortunately, the Barlow and Morgenstern *Dictionary* does not include any dynamic indications. In order to test for possible differences in dynamic levels between first and second themes, it is necessary to consult a full published score. It was decided that a subset of 50 musical works would be examined. In order to access the appropriate notated scores, we made use of the International Music Score Library Project (imslp.org). As of 2015, this repository reportedly contains some 300,000 musical scores—although this number includes many duplicate publications, as well as individual musical parts for orchestral works. The actual number of unique musical works in the IMSLP database is likely to be less than 50,000. As with the Barlow and Morgenstern, the IMSLP is hardly an unbiased random sample of Western art music. Most musicologists would regard the IMSLP as broadly biased toward piano music and orchestral works, with less
representation of chamber works. In addition, the IMSLP tends to favor classical works that are popular today.

A list of the 1063 works involving just first and second themes was randomly ordered, and we began searching for available scores on the IMSLP website. If no score was available, we proceeded to the next work on the randomly-generated list. It was determined *a priori* that only one movement would be sampled for a given work. In addition, in order to avoid a skewed sample, no more than three works were sampled by any given composer. An exception was made for Mozart: since his music was prominently represented in our random list, four works were sampled. Having located a printed score for some works, the Barlow and Morgenstern *Dictionary* was used to identify the first appearance of the first and second themes within the scores. Any associated dynamic marking was then recorded. After reaching the 182\textsuperscript{nd} item on the list, we ended sampling, having assembled 48 musical works.\footnote{An initial list of 50 works was assembled. However, it was later discovered that two of these works contained more than two musical themes. As these two works did not meet the sampling criteria, they were dropped from the list, leaving 48 musical works in the sample.}

Dynamics were coded using the traditional Italian levels: *ppp*, *pp*, *p*, *mp*, *mf*, *f*, *ff*, and *fff*. *Sforzandos*, crescendos and diminuendos were ignored. Only the basic dynamic level was coded. In the analysis, the eight levels (identified above) were treated as ordinal data. That is, *ppp* was coded as 1, *pp* as 2, *fff* as 8, etc. *A priori*, a general rule was made regarding how to code the dynamics, especially dealing with orchestral data: attention was paid only to the instruments that were playing the musical theme. For example, if the instruments that were playing the theme were marked *p* but the other accompanying instruments were marked *pp*, the dynamic level would be coded as *p*. If the dynamic level
changed during the thematic statement, or if the dynamic marking was in any way ambiguous, preference was given to the higher ordinal category (a louder dynamic level). If there was a very short period (a few notes) of disagreement within the theme, this small dissenting passage was ignored.

In orchestral works, it is common for different instruments to exhibit different concurrent dynamic levels. This might be expected where accompaniment instruments are assigned a lower dynamic level. However, surprisingly, in some scores we found that accompaniment instruments are assigned a higher dynamic level than the instruments playing the theme; in actual performance it seems unlikely that the theme would be performed objectively quieter than the accompaniment. After encountering the first instance of this anomaly, we established the following coding protocol. In circumstances where the thematic and accompaniment instruments exhibit a markedly different dynamic, we modified the coded thematic dynamic level so the dynamic levels tend to converge. Specifically, if the thematic instruments and accompaniment instruments differed by three or four ordinal categories, the ordinal coding of the thematic dynamic would be increased or decreased by one ordinal category. For example, a coding of \textit{ppp} for the theme and \textit{mp/mf} for the accompaniment would lead to a revised thematic designation of \textit{pp} (instead of \textit{ppp}). If the difference between the thematic and accompaniment dynamic levels was five or more ordinal categories, then the theme was increased or deceased by two categories. For example, if the theme was marked \textit{ppp} and the other instruments were marked \textit{f, ff, or fff}, the theme would be coded as \textit{p} (instead of \textit{ppp}). If the difference between the thematic and accompaniment instruments was in any way ambiguous (for example, if some accompanying instruments were given differing
dynamic markings from other accompanying instruments), preference was given in the
direction of the marked thematic dynamic—i.e. if the theme was marked $ppp$, with some
accompanying instruments marked $mf$ and some accompanying instruments marked $fff$,
the theme would be encoded as $pp$, only one ordinal difference from its original marking.

A further consideration in assessing dynamic level relates to the size of the
instrumentation. A theme played using $tutti$ orchestration is likely to be louder than the
same theme played by a solo instrument, even if the notated dynamic markings remain
identical. Accordingly, if a drastic change in orchestration was evident between the first
and second theme, this was taken into account in the dynamic encoding of the second
theme only. For example, if the first theme consisted of a full orchestra marked $mf$, and
the second theme consisted of only a single instrument, also marked $mf$, the second theme
dynamic code was decremented one ordinal category. In this example, the second theme
would be marked $mp$. No modifications were made to the coded dynamic level if the
changes in orchestration were deemed moderate.

It is not uncommon for notated scores to include few or no dynamic markings. If
either the first or the second theme was not marked in terms of dynamics, both the first
and second themes were encoded as the same ordinal value. For example, both were
marked $p$ if one theme was marked $p$ and the other theme was unmarked.\(^{23}\)

Finally, the change in dynamic level was determined by calculating the difference

\(^{23}\) This was done to be conservative in the statistical tests. The null hypothesis is that
there is no difference between the dynamic levels of first and second themes. By coding
themes that had unmarked dynamics as equal to the dynamic of its paired theme, the
difference value between the first and second theme would be 0. Having a difference of 0
between the dynamics of the first and second themes would make it harder to disprove
the null hypothesis of no difference.
between the coded dynamics for the first and second themes. By way of illustration, one of the randomly selected works was the fourth movement of Mozart’s *Symphony No. 12 in G, K. 110*. For the first theme (as identified by Barlow and Morgenstern), the prevailing dynamic level in the score is *forte* (coded “6”), whereas the second theme bears a dynamic marking of *piano* (coded “3”). Hence the difference in dynamic level between the two themes would be 3.

*Legato versus Smoothness*

Hypotheses 4a and 4b predict that second themes are likely to be smoother or more legato than first themes. Characterizing the degree to which a passage is “legato” is somewhat challenging. In considering possible measures, we devised two contrasting operationalizations, as described below. One method relies on a rather subjective estimate of smoothness or degree of legato by examining a variety of features in the notated scores. This labor-intensive assessment necessarily makes use of a smaller sample size. The second operationalization makes use of a simpler, more objective procedure that is readily automated. While this second method can be applied to the complete sample of 1063 paired themes, it employs a narrower notion of smoothness. Accordingly, two versions of Hypothesis 4 will be tested using contrasting methods, with the hope that the two methods will produce converging evidence.

A number of factors contribute to whether a musical passage evokes a sense of “legato.” While some of these factors reside solely in the performer’s interpretation, many can be identified in the notated score. In some cases, a composer helpfully includes pertinent performance instructions such as the terms *legato, aggitato, marcato,* or
cantabile. In other cases, the music includes explicit staccato, detaché, or accent markings. Other factors include the presence of slurs and their lengths, the appearance of many short rests between notes, sostenuto pedal markings (in the case of piano music), average note lengths in the passage, and other markings. Unfortunately, the Barlow and Morgenstern Dictionary does not include articulation, slur markings, or general performance directives.

In our first method, the degree of legato was coded on a five-point scale after an examination of the printed score. The same IMSLP sample of 48 paired themes was used for coding the articulation levels as was used for coding the dynamic levels. Specifically, thematic passages were subjectively coded as one of five possible designations, again marked on an ordinal scale: very legato (coded 1), generally legato (coded 2), balanced/unclear (coded 3), generally staccato (coded 4), and very staccato (coded 5).

Similar a priori procedures were followed during manual encoding of articulation as were used during manual encoding of dynamics. Once again, attention was paid only to the instruments that were playing the musical theme. If the articulation level changed during the theme or was ambiguous in any way, preference was given to the higher ordinal category (more staccato). If there was a very short period of disagreement (just a few notes), then the small discrepant passage was ignored. If there were grossly conflicting articulation markings in the instruments playing the thematic and accompanying material, slight modifications to the thematic articulation encoding were made. It was determined a priori that if the articulation markings differed between the two sets of instruments by three or more ordinal categories, then the thematic articulation coding would be modified by one ordinal category, either in the direction of more or less.
staccato. Unlike dynamic markings, the articulation rarely changed due to orchestration differences and was not taken into account the way it was in dynamic encodings. If either the first or second theme was unmarked in terms of articulation, it was encoded as “3” (unclear/balanced) in order to reduce possible bias. Therefore, the articulation markings were encoded conservatively. Finally, the change in articulation level was determined by calculating the difference between the coded articulation for the first and second themes.

**Measures of Rhythmic Smoothness**

Our second method for measuring “legato” is inspired by a method commonly used in speech prosody research to characterize the “smoothness” of speech. Some languages like English exhibit high inflection. That is, successive syllables tend to alternate between long and short durations, creating a sort of skipping effect. Spoken French, by contrast, exhibits a comparatively smooth or consistent rhythm in which successive syllables tend to be similar in duration.

One way to characterize the smoothness of a musical rhythm is through the pairwise variability index (PVI). This measure is commonly used by linguists to characterize the relative smoothness of syllable successions. The PVI measure calculates the difference in absolute durations of pairs of successive syllables. Syllables that exhibit contrasting durations produce large pairwise differences, whereas successive syllables that are similar in duration produce small pairwise differences. Of course, these differences are dependent on the overall speed of speaking. For slow speech the absolute differences will be larger than comparable measures when the exact same speech is

---

speeded up. In order to eliminate the effect of overall tempo, the pairwise variability is normalized, resulting in the so-called normalized pairwise variability index (nPVI).

Apart from speech, nPVI measures have been used in a number of music-related studies. Low nPVI values indicate especially smooth rhythms (e.g. *Frère Jacques*) whereas high nPVI values indicate rhythms that tend to “skip” (e.g. *English Country Garden*). Ani Patel and his colleagues have measured nPVI values for music from various European cultures, and have shown that the relative rhythmic smoothness parallel the smoothness measures for the corresponding languages spoken by the composers (Patel & Daniele, 2003; Patel, Iversen, & Rosenberg, 2006). VanHandel & Song (2010) also have shown that pairwise variability is useful for characterizing stylistic differences between individual composers. Accordingly, for the purposes of this study, we used the normalized pairwise variability index to characterize the smoothness of each of the various first and second themes. The change in rhythmic smoothness was determined by calculating the difference between the nPVI values for the first and second themes.

**Measures of Key Relatedness**

Three variants of Hypothesis 5 relate to the key-relatedness of the first and second themes. Conveniently, Barlow and Morgenstern specifically coded the key (tonic and modality) for each theme in their *Dictionary*. No effort was made to independently verify the key designations. In order to test the degree of key relatedness, some quantitative measure of key distance or similarity is needed. Traditionally, the most common metric

---

25 For more on this topic, see Patel & Daniele (2003), Huron & Ollen (2003), Patel, Iversen & Rosenberg (2006), and VanHandel & Song (2010). For a different perspective, also see London & Jones (2011).
used by theorists is the circle of fifths—where the degree of key-relatedness can be expressed as the difference in the number of accidentals in the key signature or the number of rotational steps along the circle of fifths. Accordingly, the distance between G major and D major is 1, while the distance between G major and C# major is 6. Minor keys are typically considered equivalent to their relative major keys as they share the same key signature. Arguably, one might protest that it is inappropriate to equate major and relative minor keys: surely, G major is closer to G minor than G major is to E minor. Nevertheless, we will make use of the circle of fifths as one of two key-distance measures for testing the variations of Hypothesis 5.

Both theoretical and empirical scholarship offer converging evidence for a torus-shaped representation of key relatedness (Krumhansl, 1990). Over history, a number of variants of the torus representation have been suggested. Figure 2 reproduces the torus proposed by Krumhansl (1990) on the basis of her perceptual experiments. The figure displays the torus as a flattened two-dimensional representation. One must imagine that the solid horizontal lines at the top and bottom of the figure are joined to form a horizontal tube. At the same time, the dotted vertical lines at the left and right edges are joined, completing a torus or doughnut-shaped representation.
Using the Krumhansl torus, measures of key distances can be expressed as the linear or Euclidean distances along the torus surface between the two keys. Values are scaled so that the longest distance (=1) is between a major and minor key separated by a tritone interval (i.e. C→f# or F#→c). Consequently, all key distances lie between 0 and 1.

Neither the circle of fifths nor the Krumhansl torus are claimed here as ideal measures of key relatedness. However, both might be regarded as imperfect operationalizations of the otherwise nebulous concept of key distance.

*Interval Size*

Hypothesis 6 predicts that first and second themes will exhibit different average interval sizes. We were able to calculate the average melodic interval size for each of the target first and second themes. Melodic intervals were calculated between all pairs of successive pitches. The presence of rests raises a question of whether intervals spanning rests should also be included in this calculation. For brief rests (eighth and sixteenth rests), most musicians would agree that the presence of the rest is unlikely to disrupt the
subjective experience of an interval spanning between the two notes. For longer rests (such as whole note rests), it becomes questionable whether the antecedent and consequent pitches evoke the subjective experience of an interval. Since the Barlow and Morgenstern *Dictionary* explicitly encodes passages deemed “themes,” we assumed that all of the notes in the theme would be part of a single musical idea, and so all intervals were calculated, including all intervals spanning rests, without regard to the duration of the rest. The difference in average interval size between first and second themes was calculated.

Results

*Modality*

In the case of Hypothesis 1, we predicted that second themes are more likely to be in the minor mode. Using McNemar’s test with a binomial distribution, first themes (326/1063, 30.67%) were more likely to be in the minor mode than were second themes (183/1063, 17.21%), $p < 0.000001$. Examining the observed proportions, it was also found that first themes in the minor key were more likely to have a second theme also in minor (76/326, 23.3%) than first themes in the major key (107/737, 14.5%). Similarly, first themes in the major key were more likely to have second themes in the major key (630/737, 85.5%) than were first themes in the minor key (250/326, 76.7%). As can be seen in Figure 3, there appears to be less use of the minor mode in second as opposed to first themes. This association is not consistent with the hypothesis and has interesting implications, to be discussed below.
Figure 3. Second themes are less likely to be in the minor mode than are first themes. Numbers represent the proportion of themes in the minor mode, with standard error bars shown.

Since these results were not expected, we wondered if works in sonata-allegro form could have had significant influence on the works in other structural forms. Therefore, we wanted to remove the influence of musical pieces in sonata form and run the same statistical test. Instead of analyzing the structure of all 1063 works in our sample, we decided to operationalize a method of roughly identifying works that were likely in sonata form. We deemed that any first movement of a musical work called Sonata, Concerto, Symphony, Trio/Quartet/Quintet/Sextet/Octet was likely to be in sonata form; therefore we removed all of the pieces that met this criterion from our sample. With this new sample (n = 873), we performed another McNemar’s test with a binominal distribution. The results from the post hoc test were the same as the results of the initial test with the complete data set, p < 0.05.
Tempo

In the case of Hypothesis 2, we predicted that second themes are more likely to exhibit a slower pace as measured using a “pace” value formula. Recall that larger “pace” numbers are associated with a greater density of short-duration notes. The difference between the two pace measures was determined by subtracting the second theme pace value from the first theme pace value. Using the Wilcoxon Signed-Ranks Test, second themes (mean pace value = 7.647, sd = 4.00) were more likely to exhibit slower pace values than were first themes (mean pace value = 9.01, sd = 7.49), where longer note durations produce lower pace values, Z = -9.09, p < 0.000001. The average difference was found to be 1.136, consistent with the hypothesis that first themes tend to be faster than second themes. The average pace values for first and second themes are shown graphically in Figure 4.

Figure 4. Second themes exhibit a slower pace value than do first themes. Numbers represent the average pace value for first and second themes, with standard error bars shown.
Dynamic Level

Recall from the methods section that we sampled 48 pairs of dynamic levels for first and second themes. Recall also that the dynamic levels were coded as ordinal values from 1 (representing ppp) to 8 (representing fff). The difference between the two dynamic levels was determined by subtracting the second theme dynamic level from the first theme dynamic level. Using the Wilcoxon Signed-Ranks Test, second themes (mean dynamic level = 4.04, sd = 1.62) were more likely to exhibit quieter dynamic levels than were first themes (mean dynamic level = 4.67, sd = 1.72), where higher numbers correspond to louder dynamic levels, Z = -2.019, p = 0.044. The average difference was found to be 0.626, which is consistent with the hypothesis that second themes tend to be quieter than first themes. The results are pictured graphically in Figure 5.

![Figure 5. Second themes exhibit quieter dynamic levels than do first themes. Numbers represent the average dynamic level of first and second themes, with standard error bars shown.](image-url)
Legato

With regard to articulation, recall that we made use of two methods. In the case of the IMSLP sample, we can contrast the ordinal numerical ratings for the contrasting first and second themes. The difference between the two articulation ratings was determined by subtracting the second theme’s articulation rating from the first theme’s articulation rating. Using the Wilcoxon Signed-Ranks Test, second themes (mean articulation level = 2.79, sd = 1.30) were likely to exhibit more legato markings than were first themes (mean dynamic level = 3.23, sd = 1.43), where higher numbers correspond to less legato/more staccato markings, \( Z = -2.135, p = 0.033 \). The average difference was found to be 0.438, which is consistent with the idea that second themes tend to be more legato than first themes. These results are shown in Figure 6.

![Figure 6. Second themes are more legato than are first themes. Numbers represent the average articulation level of first and second themes, with standard error bars shown.](image)

45
Smoothness

As an additional test of this hypothesis, we also compared the normalized pairwise variability index for the first and second theme rhythms. Recall that our hypothesis specifically relates to the degree to which rhythms are isochronous. The difference between the two nPVI values was determined by subtracting the second theme nPVI values from the corresponding first theme nPVI values. Using the Wilcoxon Signed-Ranks Test, second themes (mean nPVI = 41.40, sd = 22.71) were not likely to exhibit more isochronous rhythms than were first themes (mean nPVI = 42.27, sd = 25.72), where lower numbers correspond to more smooth rhythms, $Z = -0.390, p = 0.697$. The pertinent summary data are shown graphically in Figure 7.

![Figure 7](image)

Figure 7. First and second themes do not differ in terms of nPVI values. Numbers represent the average nPVI values of first and second themes, with standard error bars shown.

In light of this failure, we might consider a further conjecture regarding the rhythmic smoothness of first and second themes. It may be that the themes are simply
contrasting with one another, without regard to whether first or second themes exhibit greater rhythmic smoothness. In order to test this post hoc hypothesis, a bootstrap simulation was conducted in which first and second themes were systematically permuted with each other. That is, every first theme was paired with every other second theme in our data set, resulting in over a million thematic pairings. The absolute differences in nPVI values were then determined for both the randomized thematic pairings and the actual pairings. The average actual pairings differed by 24.9 nPVI units, whereas the average random pairings differed by 26.9 units—suggesting that difference between actual thematic pairings are generally smaller than random thematic pairings. In order to determine whether this difference is statistically significant, a series of 10,000 random samples was drawn from the complete permuted pairings and means were calculated for each sample. Each random sample was equivalent in size to the total number of actual pairings—allowing us to directly calculate the probability of drawing samples with an average value of 24.9 or lower. None of the 10,000 random samples provided a lower mean than the actual-pairings mean, hence \( p < 0.0001 \). Contrary to the post hoc hypothesis, first and second themes are not more contrasting than would be expected by chance with regard to pairwise-variability in rhythm. Instead, first and second themes tend to more closely resemble each other in terms of the degree of rhythmic smoothness.

**Key Relationships**

As noted earlier, the relationship between the keys for the first and second theme-pairs was measured using two methods: the circle-of-fifths key distance (measured in numbers of accidentals) and the Krumhansl torus distance. The first hypothesis about
keys (Hypothesis 5a) stated that first and second themes were likely to be in different keys. Comparing the keys of first and second themes using a one-sample binomial test with a test-proportion of 0.50, it was found that first and second themes were, in fact, more likely to be in different keys (proportion = 0.67, 717/1063), \( p < 0.000001 \).

The second hypothesis about key relationships between first and second themes stated that, compared to their paired first theme, second themes were more likely to be in a closely related key than they were to be in a distantly related key. Using the circle-of-fifths measure, across the complete sample of 1063 thematic pairs, the average key distance was 1.6764 accidentals (sd = 2.1834). This key-distance measure alone is not sufficient to determine whether themes tend to be closely related. Keys like G major and D minor are more common in music than A# major or F# minor. Consequently, we need to consider the typical key distance between unrelated themes. Accordingly, rather than simply treating each key as equally probable, we created a control comparison by randomly assigning pairs of first and second themes, and measuring the key-distance between these random pairs. The scrambled thematic pairs resulted in an average distance of 3.0517 (sd = 2.2508). Using a t-test, it was found that the distance between first and second themes (mean = 1.676, sd = 2.183), on average, are smaller in size than are scrambled thematic pairs (mean = 3.052, sd = 2.251), \( t(1161) = -6.01 \), \( p = < 0.00001 \).

Using the Krumhansl torus distance measures, the average distance between the first and second themes was determined to be 0.2824 (sd = 0.2342). For the scrambled pairings of first and second themes, the corresponding distances had an average of 0.5011 (sd = 0.2307). Again, we see a smaller distance between the keys of first and second themes than in the keys of the scrambled thematic pairs.
Whether measured using the circle-of-fifths or using the Krumhansl key torus, the keys of the first and second themes are clearly closely related. A follow-up classification was carried out in order to identify the most common types of key relations, relating to Hypothesis 5c. The 15 most common key relationships are reported in Table 2. The most common relation (25.8%) has both themes sharing the same key. When relative key relationships are included, nearly 40 percent of theme pairs share the same key signature. The second most common relationship involves modulation to the dominant (18.25%). Themes that differ by one sharp or flat represent 24 percent of theme pairs. Including enharmonic spellings, nearly two-thirds (61.3%) of theme pairs either share the same key signature or differ by one accidental. Other key relationships—notably Neapolitan and tritone relationships—occurred with frequencies less than 1 percent.

<table>
<thead>
<tr>
<th>Key relationship</th>
<th>Example</th>
<th>Frequency (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same key</td>
<td>C→C (c→c)</td>
<td>25.78</td>
</tr>
<tr>
<td>To dominant</td>
<td>C→G (c→G)</td>
<td>18.25</td>
</tr>
<tr>
<td>To relative major</td>
<td>c→Eb</td>
<td>8.75</td>
</tr>
<tr>
<td>To parallel major</td>
<td>c→C</td>
<td>6.30</td>
</tr>
<tr>
<td>To subdominant</td>
<td>C→F</td>
<td>5.83</td>
</tr>
<tr>
<td>To parallel minor</td>
<td>C→c</td>
<td>4.14</td>
</tr>
<tr>
<td>To relative minor</td>
<td>C→a</td>
<td>2.73</td>
</tr>
<tr>
<td>To subtonic major</td>
<td>c→Bb</td>
<td>2.63</td>
</tr>
<tr>
<td>To lowered submediant</td>
<td>C→Ab</td>
<td>2.63</td>
</tr>
<tr>
<td>To lowered mediant</td>
<td>C→Eb</td>
<td>2.36</td>
</tr>
<tr>
<td>To mediant</td>
<td>C→E</td>
<td>2.26</td>
</tr>
<tr>
<td>To submediant</td>
<td>C→A</td>
<td>2.07</td>
</tr>
<tr>
<td>To mediant major</td>
<td>c→E</td>
<td>1.69</td>
</tr>
<tr>
<td>To supertonic</td>
<td>C→D</td>
<td>1.13</td>
</tr>
<tr>
<td>To subtonic</td>
<td>C→Bb</td>
<td>1.13</td>
</tr>
</tbody>
</table>

Table 2. Most frequent key relationships between first and second thematic pairs.
Average Interval Size

Apart from the *a priori* proposed hypotheses, we also intended to explore possible differences in interval size between first and second themes. Recall that we offered no specific hypothesis as to the direction of change. Once again, we calculated the difference between the average interval sizes of the first and second themes. Using the Wilcoxon Signed-Ranks Test, second themes (mean interval size in semitones = 2.43, sd = 0.903) were more likely to exhibit smaller interval sizes than were first themes (mean interval size in semitones = 2.56, sd = 1.08), where higher numbers correspond to larger average intervals, $Z = -2.802$, $p = 0.005$. However, note that the difference between the average interval size in first and second themes is small. The average melodic interval sizes (in semitones) for first and second themes are shown in Figure 8.

![Figure 8](image)

Figure 8. Second themes utilize slightly smaller intervals than do first themes. Numbers represent the average interval sizes of first and second themes, with standard error bars shown.
Stylistic Changes Over Time

As mentioned earlier, music theorists have suggested a number of ways in which first and second themes have varied in their musical properties over time. *Post hoc*, we carried out a series of statistical tests in order to determine if some of these trends could be seen in our sample.

Tracing historical changes presupposes we can date each of the musical works in our sample. Although the *Dictionary of Musical Themes* provides the composers’ biographical dates, it does not indicate the date of composition for the individual themes. However, the electronic version of this database has been augmented to include dates for 7,984 of the 9,788 themes (80%). Where known, these dates specify the year of composition. However, in many cases, only the date of publication is known. In this study, we made use of the date of composition (where known), and otherwise made use of the date of publication. In some cases, publication of a musical work is posthumous. In these cases, we used the death-year of the composer rather than the date of publication. In many cases, the date of composition or publication is given as a range (such as 1746-1749). In order to simplify processing, in these cases, we used the first date in the range.

In presenting historical data, it is convenient to partition the information by historical period or era. One might elect to present the data according to century, however there is some merit in presenting the information by stylistic period. Acknowledging that dates for historical periods are rather arbitrary and that the notion of “period” is highly contested, we nevertheless elected to partition the data into suitable eras. For this purpose we made use of the dates suggested by Naxos Records for various music-stylistic periods. We first determined the distribution of works from our sample across historical time. In
our main sample (n = 1063) we found that only two works were from the period 1400-1599 (nominally “Renaissance”), 43 were from the period 1600-1749 (nominally “Baroque”), 351 were from the period 1750-1829 (“Classical”), 103 were from the nominal Early Romantic period 1830-1859 (“Early Romantic”), 425 were from the period 1860-1919 (“Romantic”), and 139 were from the period 1920-1959 (“Post-Great War” period). In our smaller sample (n = 48) used for articulation and dynamic markings, 0 were Renaissance works, 1 was a Baroque work, 9 were Classical works, 4 were Early Romantic works, 23 were Romantic works, and 11 were Post-Great War works. Similar biases can be seen across both samples.

Using the same nonparametric tests as in the main study, we separately tested whether first and second themes differed in mode, pace value, dynamics, articulation, npV1, key distance, and average interval size in each time period. Renaissance music was excluded due to its small sample size. Thirty post hoc tests were conducted, so all significance values have been adjusted using the Bonferroni correction.

Two observations were made regarding differing stylistic treatment of first and second themes across musical periods. As noted earlier, across all periods, first themes are more likely to appear in minor than second themes. However, as evident in Figure 9, both first and second themes are more likely to make use of the minor mode during the Early Romantic and Romantic periods compared with earlier and later periods. This result is consistent with the results of Horn and Huron (2015).

As already noted, second themes tend to be slower paced than first themes. However, Figure 9 shows that the difference is more marked in the Early Romantic and Romantic periods. It bears emphasizing that these differences represent averages and
thus, one should use caution when interpreting these results. For example, the results do not necessarily mean that the pace values in the Early Romantic and Romantic periods were more contrasting than in other periods. Theoretically, greater contrasts could exist in the Baroque period, except that more works exhibited faster paced second themes than first themes. Extreme values can lead to moderate averages.

![Figure 9. Comparison of modality (top graph) and pace value (bottom graph) in first and second themes across musical periods. Significant differences between first and second themes are marked with an asterisk (*) and standard error bars are shown.]

Themes With an Introduction

Many musical works begin with an introductory passage; perhaps the best-known example is sonata form musical works that include a slow introduction. This raises the question of whether the relationship between first and second themes might change when
there is an additional introductory theme. Our original sample consisted of musical works
that contained precisely one first and one second theme—according to Barlow and
Morgenstern’s analyses. However, a small sample of 18 musical works containing an
introductory theme and one first and one second theme was also available for study. We
performed the same statistical tests in this supplementary sample that we did in the main
study; however, we did not test for articulation and dynamics, as there were no works that
had been coded for articulation and dynamics that also included an introduction. The
results from these tests are reported in Table 3. Possibly due to the small sample size,
none of the statistics reach significance. However, all of the results are skewed in the
predicted direction, suggesting that there may be little difference between two-themed
works with and without an introductory theme.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>A Theme</th>
<th>B Theme</th>
<th>Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent in Minor</td>
<td>50% (9/18)</td>
<td>22.22% (4/18)</td>
<td>X²(1)= 2.286</td>
<td>0.131</td>
</tr>
<tr>
<td>Pace</td>
<td>6.620 (sd=2.03)</td>
<td>6.146 (sd=2.38)</td>
<td>Z = -1.067</td>
<td>0.286</td>
</tr>
<tr>
<td>Interval Size</td>
<td>2.37 (sd=0.62)</td>
<td>2.37 (sd=0.88)</td>
<td>Z= -0.414</td>
<td>0.679</td>
</tr>
</tbody>
</table>

Table 3. Results of statistical analyses for works exhibiting one introductory theme, one
first theme, and one second theme. Although none of the results are statistically
significant, the direction of all of the results trend in the same direction as the musical
samples with no thematic introduction.

Discussion

The goal of this corpus study was to empirically test whether common theoretical
ideas regarding first and second themes can be observed in a large sample of music.
Although much of the historical discussion regarding thematic content focuses on sonata-
allegro form expositions, the samples we tested included a broad range of musical forms,
genres, and ensembles. Our observations regarding the pace values, dynamic levels, and
articulations are consistent with traditional textbook descriptions that first themes are *energetic* in nature and second themes are *lyrical* in nature, given our operationalizations. Although there are many exceptions to this generalization, our results are consistent with the practice of contrasting expressive content in first and second themes.

One exception to the energetic/lyrical polarity is evident in the choice of modes. In keeping with the idea of energetic first themes and lyrical second themes, we conjectured that second themes would be more likely to be in the minor mode. However, the results showed a reverse tendency. Although most of our sampled works are not sonata expositions, this reverse relationship is nevertheless in keeping with traditional sonata-allegro theory (Hepokoski & Darcy, 2006). That is, when first themes are in the minor mode, second themes tend to employ a corresponding major mode—either relative or parallel; when first themes are in the major mode, second themes also tend to favor the major mode—principally the dominant. This finding implies that key relationships between first and second themes may be dominated by concerns regarding modulation rather than the idea that the minor mode might be more suitable for lyrical expression. This result is consistent with Rosen’s (1971) view of the difference between first and second themes.

Given the findings in this study, there seems to be an intriguing relationship between characteristics of second musical themes and the characteristics of sad music. The characterization of themes that we operationalized and tested was based on music theorists who described second themes as *lyrical* and *gentle*. With the exception of mode, second themes exhibit many characterizations of sad music—they are slow, legato, quiet,
and have smaller pitch movements (average interval size). The fact that second themes are similar and yet distinct from characterizations of sad music is interesting. It could be that the higher proportion of second themes in the major mode is what distinguishes sad music from “tender-lyrical” music (Horn and Huron, 2015). Future research in this direction is warranted.

With regard to the rhythmic or articulatory smoothness, our research found that first and second themes do not differ in terms of their rhythmic nPVI values, while they do differ with respect to articulation. We additionally saw that first and second themes have more similar nPVI values than do randomly paired first and second themes. These contrasting findings suggest that more legato second themes are created principally through notated articulation, rather than by composing smoother duration values.

The aims of the corpus study were twofold: we asked (1) if we could observe the contrasts between first and second themes described in literature from music theory and (2) if we could observe changes in the characteristics of first and second themes over different stylistic musical periods. Neither of these goals gives any indication if listeners can perceive these structural differences in musical excerpts. A modest attempt at understanding the relationship between these structural musical features, the musical score, and recorded audio is the subject of Chapter 4.

For more information about characteristics of sad music, see the following sources: Paul and Huron, 2010; Turner and Huron, 2008; Post and Huron, 2009; Huron, 2008; and Schutz, Huron, Keeton, and Lower, 2008.
Chapter 4: Perception of First and Second Themes from Musical Scores and Sound Recordings

As we have seen in the first three chapters, there have long been descriptive differences between first and second musical themes. In the corpus study, we examined structural features of a large body of musical themes, written over a period of 300 years. The results were consistent with the hypothesis that, compared to first themes, second themes are less likely to be in the minor mode, and are more likely to be legato, make use of a slower pace, use smaller intervals, and utilize a quieter dynamic level. However, it is not clear whether people can perceive these differences in real music or whether they go largely unnoticed. Therefore, the goal of the next study was to determine whether participants could differentiate between a pair of first and second musical themes using features of the musical construction.

Recall that we operationalized descriptive characteristics of musical themes into six measurable criteria. We hypothesized that compared to first themes, second themes would be more likely to be in the minor mode, exhibit a slower tempo or pace, have a quieter dynamic level, have more legato articulations, exhibit smoother rhythms, and differ in average interval size. Each of these individual features has been shown to be
perceptible to listeners.\textsuperscript{27} Of course, real music is complex, so perceptual studies that isolate one or two specific features are frequently low in ecological validity. In an attempt to preserve some degree of ecological validity, the current study endeavors to examine multiple features at once.

A few research studies have shown that participants are able to perceive differences in multiple surface level features of musical samples, similar to the current task at hand. For example, McAdams \textit{et al.} (2004) showed that structural features of the music, such as melody, rhythm, and articulation, can be used to categorize musical stimuli. Keller and Schubert (2011) demonstrated that people can use surface level features to group musical stimuli according to affect. A few studies have also indicated that performance features, such as tempo and dynamics, can alter participants’ perceptual judgments about musical stimuli. Repp (1998) showed that differences in expressive timing can affect perceptual judgments in musical samples. In terms of dynamics, Kamenetsky, Hill, and Trehub (1997) found that dynamic variance can lead to judgments of higher emotional expressiveness in MIDI recordings.

The current study aimed to determine whether people can differentiate between a pair of musical themes using perceptual factors. The musical features of interest were the structural features examined in the corpus study: average interval size, pace value, rhythmic smoothness, articulation, and dynamics. The study focused on the relationship

\footnotesize{
}
between first and second themes from the same musical work and consisted of music composed in the eighteenth through twentieth centuries.

Hypotheses

The question at hand is to determine which musical features listeners use in order to distinguish between first and second musical themes. In daily life, this task most directly applies to how listeners hear musical performances. However, the way in which a performer brings a piece of music to life can exaggerate, eliminate, or alter the musical criteria written in the score. For example, existing knowledge of sonata form could lead a performer to exaggerate the “energetic” first themes and the “lyrical” second themes. Alternatively, performers may wish to convey information that they believe is lost in the musical score, based on extra-musical knowledge. They may also wish to provide a new interpretation of a well-known piece. Playing a typically “graceful” second theme in a sardonic manner will obscure the musical content provided in the score. Therefore, any hypothesis must be made in regard to music heard without interpretive performance nuance. This leads to the first hypothesis:

H1. Participants can categorize first and second themes when listening to MIDI sound recordings.

MIDI recordings take score-based data and produce an auditory recording without any performance nuances. However, with the advent of recent technology, it is possible to have MIDI recordings that differ in features like tempo, dynamics, and articulation. For the purposes of this study, we were mostly interested in how variations in articulation and
dynamics affected performance on the discrimination task.\textsuperscript{28} Since the inclusion of articulation and dynamics is likely to help in the discrimination between first and second themes, a corollary to the first hypothesis was made:

\textbf{H1b.} People are better at categorizing themes from sound recordings when dynamics and articulation are included than from sound recordings without these features.

An even more reductive approach can be taken when considering the written musical score. Discriminating between first and second themes from a score requires a different set of skills than does an aural analysis. However, the notated music likely does contain information that can allow a person to distinguish between two written themes. For example, the corpus study presented in Chapter 3 was conducted only in reference to the notated music. The interest in musical scores leads to a second hypothesis and its corollary:

\textbf{H2.} Participants are able to distinguish first and second musical themes by looking at the information contained in the musical score.

\textbf{H2b.} People should perform better on the task when the provided scores include dynamics and articulation than when the given scores do not include this information.

Since the analysis of written scores requires familiarization with musical notation, and therefore some training, it could be expected that people with more musical training would do considerably better than people with less musical training in the score-based perception task. Although musical training should also facilitate performance on the

\textsuperscript{28} These two features, articulation and dynamics, will be manipulated in the main study.
recording-based perception task, the difference in performance between people of various levels of musical training may be smaller in the case of the recording-based task. No hypothesis was made about whether people with a high level of musical training would be better at the score-based or recording-based perception tasks, since, to some degree, they are different skills.

The structural features operationalized and measured in the corpus study have been shown to differ between first and second themes (see Table 4). Accordingly, when differentiating between two themes, musicians should be able to use these criteria to make their judgments. Additional criteria may be used to make these distinctions, as well, but no \textit{a priori} hypotheses were made regarding other criteria. Importantly, the consideration of harmony and phrase-structure were not considered as \textit{a priori} hypotheses. We saw in Chapter 2 that harmonic movement, cadence type, and the organization of the phrases (tight-knit or loosely organized) are considered by some theorists to be important features in categorizing theme types (Caplin, 1998). These features were not considered in the corpus study because they are not available in the Barlow and Morgenstern \textit{Dictionary of Musical Themes}. In order to compare the perceptual and corpus studies, it was decided to retain the same criteria in the perceptual study as was tested in the preceding corpus study. Harmonic implications will be discussed more at the end of the chapter and in Appendix C.
Table 4. Criteria from the corpus study, which were hypothesized to differ between first and second themes. These criteria are expected to play a role in judgments about which themes appear first or second in a piece of music. The italicized items, Mode and Key, were not considered in the current study.

<table>
<thead>
<tr>
<th>Score-Based Criteria</th>
<th>Performance Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pace Value (proxy for tempo)</td>
<td>Articulation (legato, staccato)</td>
</tr>
<tr>
<td>Average Interval Size</td>
<td>Dynamics</td>
</tr>
<tr>
<td>Rhythmic Smoothness (nPVI)</td>
<td></td>
</tr>
<tr>
<td><strong>Mode (major or minor)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Key</strong></td>
<td></td>
</tr>
</tbody>
</table>

Sample

For the purposes of this study, an initial sample of 166 piano works was taken from Sam Barlow and Harold Morgenstern’s *Dictionary of Musical Themes* (1948). The works were a subset of the main sample used in the corpus study. Therefore, the selected works contained only two themes. This means that pieces with an introductory theme or that had more than two themes were excluded. Approximately half of the sample consisted of works in sonata form, but the other half of the sample represented music written in other structural forms. The piano works in the sample spanned compositional dates from 1777 to 1939 and were biased towards Classical and Romantic music. The final sample of musical works is listed in Appendix B. Using the Humdrum...
Toolkit (Huron, 1993), the normalized pairwise variability index (nPVI), average pace value, average interval size, mode, and key were estimated. In addition to these measures, manual coding of articulation and dynamics was carried out using an ordinal classification for all 166 piano works. Ordinal levels were determined in the same way as they were in the corpus study.

The comparison of thematic pairs from a single musical work could lead musically-trained participants to respond in one of a few ways. The first outcome is that participants distinguish between the musical cues written by the composer, using one or more characteristics of the music in order to make their judgment. Participants may use subtle criteria, such as variations in interval sizes, articulation, and dynamics to distinguish between two themes. Another possibility is that participants use a heuristic to categorize first and second themes. A common heuristic would be that a major theme is a second theme, whereas its paired minor theme is a first theme, given conventions about primary and secondary themes in sonata form. Attempts were made to circumvent the use of such heuristics. Therefore, the sample consisted of first and second themes that are in the same mode. Additionally, all of the samples were transposed to have a tonal center of C: musical pairs that were in a major mode were transposed to C Major and musical pairs that were in a minor mode were transposed to c minor. The transposed samples prevented key relationships (such as I – V or i – III) from affecting the discrimination between first and second themes. Therefore, the categorization of a pair of musical themes must be done with regard to more subtle musical criteria, instead of only using basic knowledge about music theory.
The initial sample of 166 pairs of musical themes was evaluated for how different their first and second themes were in terms of their dynamic and articulation levels. Based on this comparison, the musical pairs were grouped into one of four categories, as shown in Table 5. The corpus study was consistent with the idea that first themes tend to be louder and more staccato than are secondary themes. Thus, musical pairs that follow this pattern of articulation and dynamics were deemed “Expected,” as they followed the trends exhibited by a large body of music. Contrastingly, musical pairs that followed the reverse trend, with first themes that were quieter and more legato than the paired second themes, were deemed “Opposite,” as they exhibited a pattern that was opposite of what was expected, given the findings in the corpus study. The two other groups referred to the remaining possible relationships between the dynamics and articulation of first and second themes. Pairs of themes where the first themes were louder and more legato than were the second themes were assigned to the “Contrasting Articulation” category, as they followed the expected pattern of dynamics, but the opposite pattern of articulation. Musical pairs where the first themes were quieter and more staccato than the second themes belonged to the “Contrasting Dynamics” category, as the pattern of articulation was expected, but the pattern of dynamics was not.

The final sample of musical stimuli (n = 22) was chosen so that these four categories were balanced, with approximately the same number of musical pairs in each of the four categories.

Using musical thematic pairs that may only sometimes conform to prior expectations has an interesting implication for the current perceptual study. If we accept the null hypothesis as true – that people can use specified criteria to make distinctions

---

31 These two features, articulation and dynamics, are the features that will be manipulated in the main study.
between first and second themes, consistent with generalizations given in music theory—
we would expect that themes that are softer, more legato, have smaller intervals, and have a slower pace value than their thematic pair will be accepted as “second themes,” regardless of their actual presented order in the music. This means that discriminations about thematic pairs that are in the “Expected” category should be answered correctly and discriminations about thematic pairs that are in the “Opposite” category should be answered incorrectly.

In this study, the “Expected” and “Opposite” categories were defined exclusively according to the dynamics and articulation of the thematic pairs. The other structural features—interval size, durational pace, and nPVI—were not considered in the categorizing the thematic pairs. Altering a theme’s articulation and dynamics will not alter the pitch or rhythm of the melodies. However, changes in nPVI, average interval size, and durational pace would all necessitate some alteration to the pitch and/or rhythm. In order to facilitate comparison between themes in the corpus and perceptual studies, it was decided *a priori* to leave the thematic melodies as they were originally written by the composers.

<table>
<thead>
<tr>
<th>Expected (n = 6)</th>
<th>Opposite (n = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(First themes louder, more staccato)</td>
<td>(First themes quieter, more legato)</td>
</tr>
<tr>
<td><strong>Contrasting Articulation (n = 5)</strong></td>
<td><strong>Contrasting Dynamics (n = 5)</strong></td>
</tr>
<tr>
<td>(First themes louder, more legato)</td>
<td>(First themes quieter, more staccato)</td>
</tr>
</tbody>
</table>

Table 5. Categorization of musical stimuli based on their articulation and dynamic levels. The categories are assigned with respect to how similar they are to the generalizations seen in musical pairs in the corpus study.
Method

The study consisted of two parts, both based on a two-alternative forced-choice task. The first part asked participants to examine the musical score of a pair of musical themes that came from the same work; they were asked to identify which theme they believed comes first in the music. The second part requested the same participants to listen to recordings of the same musical thematic pairs and to identify which theme was presented first in the music. Both the score-based task and the recording-based task consisted of two conditions: A) musical samples with dynamic and articulative information and B) musical samples without dynamic and articulative information.

Score-Based Perception

One of the hypotheses of this study is that participants are able to categorize first and second musical themes by looking at the information contained in the musical score. The task used to test this hypothesis required participants to distinguish between a pair of themes that came from the same musical work. In order to avoid variability between different published musical editions, the first and second themes were encoded into the musical software Finale 2014. The themes were written as they first appear in the musical score. All external cues were eliminated from the Finale notation. This included types of information such as measure numbers, tempo markings, and headers like “Minuet” or “Trio.” Performance indications such as “animato,” “dolce,” and “furioso” were also eliminated.

32 It should be noted that the first occurrence of the theme is not necessarily the most iconic rendition of the theme. A good example is the fourth movement of Beethoven’s Symphony 9, where the original statement of the first theme begins with a half note (F# G A A ….), rather than with the two quarter notes most people recognize as the theme (F# F# G A A ….).
removed from the thematic samples. In this study, accompaniment of the theme was included, allowing for some harmonic cues to be available to the participants. A discussion of this decision will occur at the end of the chapter.

The hypothesis also specified that participants should perform better on this task when they are given access to articulation and dynamic information. Therefore, two versions of the musical samples were encoded into Finale 2014. The first version consisted of the musical notes, dynamic markings, and articulation markings; this version was deemed the “With Dynamics and Articulation (With DA)” sample. A second version removed the dynamic and articulation markings and was deemed the “Without Dynamics and Articulation (Without DA)” sample. Therefore, in the “Without DA” sample, participants could only use certain features of the musical structure (such as nPVI, average interval size, and pace value) to make the determination of “first” and “second” themes. An example of each of the two conditions is shown in Figure 10.

Participants were presented with two themes from a musical work on a single screen. They were asked to categorize the pair of themes by indicating which theme they believed came first in the music and which theme they believed to appear second in the music. Additionally, they were asked to rate the confidence of their response on a 10-point sliding scale. The presented order of musical works was randomized, as was the order of the first and second themes for each musical work. Two samples were chosen to be used twice as an indication of test-retest reliability.

In an attempt to circumvent demand characteristics, it was decided that each participant would not be presented with both the “With DA” and “Without DA” versions of each musical pair. Instead, the study consisted of two experimental groups, utilizing a
between-subjects design. The groups were counterbalanced with regard to dynamic and articulation condition. In Participant Group 1, 11 of the musical pairs were given “With DA” and 11 musical pairs were shown “Without DA.” In Participant Group 2, these conditions were reversed, so that the musical pairs that were in the “With DA” condition in Participant Group 1 were the musical pairs that were in the “Without DA” condition in Participant Group 2. The “With DA” and “Without DA” samples were balanced across the four categories of musical stimuli (Expected, Opposite, Conflicting Articulation, Conflicting Dynamics), resulting in the arrangement in Table 6. Therefore, every musical thematic pair was used in the experiment with dynamics and articulation and without dynamics and articulation, but each participant only saw each thematic pair in one of the two conditions. This design allows overall performance of the music categorization to be compared across the two conditions. Presumably, people will perform better in the “With DA” condition than in the “Without DA” condition.

Figure 10. The same musical score shown in the “Without Dynamics and Articulation” condition (top line) and in the “With Dynamics and Articulation” condition (bottom line). The music is the first theme of Dvorak’s *Waltzes, Op. 54, No. 3*. 

68
<table>
<thead>
<tr>
<th>Musical Category</th>
<th>Condition</th>
<th>Number of Music Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected</td>
<td>With DA</td>
<td>3</td>
</tr>
<tr>
<td>Opposite</td>
<td>With DA</td>
<td>3</td>
</tr>
<tr>
<td>Conflicting Articulation</td>
<td>With DA</td>
<td>3</td>
</tr>
<tr>
<td>Conflicting Dynamics</td>
<td>With DA</td>
<td>2 (or 3)</td>
</tr>
<tr>
<td>Expected</td>
<td>Without DA</td>
<td>3</td>
</tr>
<tr>
<td>Opposite</td>
<td>Without DA</td>
<td>3</td>
</tr>
<tr>
<td>Conflicting Articulation</td>
<td>Without DA</td>
<td>2 (or 3)</td>
</tr>
<tr>
<td>Conflicting Dynamics</td>
<td>Without DA</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 6. Arrangement of musical stimuli (n = 22) as a cross of musical category and condition. This design was counterbalanced across two participant groups, so the musical pairs in the “With DA” and “Without DA” conditions were switched between the two participant groups.

Recording-Based Perception

The second major hypothesis is that people should be able to discriminate between first and second themes from audio recordings. The corollary hypothesis is that recordings including dynamics and articulation should result in better overall performance than recordings without these two features. Therefore, similar to the score-based perception task, the recording-based task involved two conditions: one with the inclusion of articulation and dynamic information (“With DA”) and one without these features (“Without DA”).

In order to compare a single participant’s performance on the score-based and recording-based perceptual tasks, a within-subjects design was maintained across the two major parts of the study. This means that for a given musical sample, the condition (either “With DA” or “Without DA”) was maintained across both parts. If a participant saw the score to a Mozart sonata with articulation and dynamics and saw the score to a Beethoven sonata without articulation and dynamics, he or she would hear the recording of the
Mozart sonata with articulation and dynamics and hear the Beethoven recording without articulation and dynamics. The musical recordings were therefore counterbalanced in the same manner as were the musical scores. In summary, every musical recording in the experiment was heard with and without dynamics and articulation, but each participant only heard one of the two versions of the recordings.

To eliminate bias from real human performance, each recording of the piano music was made from the Finale notation file, using an automated setting called “Human Playback.” The tempo chosen was the one indicated by the score and by experience of the author, who was familiar with all of the piano music in the sample. Two separate recordings of the piano works were obtained from Finale: a control condition (akin to “Without DA”) and a performance condition (akin to “With DA”). In the performance condition, the audio file included articulation and dynamic markings. The Human Playback setting in Finale 2014 allows these settings to be expressed in audio recordings. The control condition also used audio renderings of the Finale scores, but without any audible articulation or dynamic effects.

The design of the recording-based task was the same as the design of the score-based task. Participants heard audio files of the first and second themes from the twenty-two piano works. They were told they could listen to each theme as many times as they wished. Their task was to indicate which theme came first in the music and which theme appeared second in the music. They were also asked to rate how confident they were about their response using a 10-point sliding scale. The presented order of musical works was randomized, as was the order of the first and second themes for each musical work. Two samples were chosen to be used twice as an indication of test-retest reliability. In
addition to whether or not participants could complete the task, the difference in overall performance between “With DA” and “Without DA” conditions was a point of interest.

Protocol

Participants completed the two tasks online, via the Qualtrics Research Suite. In order to account for individual variability, the study was created with a within-subjects design, so participants completed the recording- and score-based tasks with the same conditions of the musical pairs (either with or without dynamic and articulation information). The participants were mainly second year music theory students at the Ohio State University. Many first and second year theory students have weak audiation abilities. Namely, students are not able to look at a musical score and hear the music in their heads. The weakness of students’ audiation abilities has been noted in my own experiences teaching Theory I and II, as well as in the experiences of other music theory instructors. Therefore, it is unlikely that these participants would remember the musical score (and his or her response to the discrimination) and be able to connect it to a recording in the later section. Conversely, it is more likely that a participant would remember the sound recording and be able to connect it to a musical score seen later, as this is a common task in music theory training. Because of this fact, all participants completed the score-based perception task before they completed the recording-based task. In both the recording-based and score-based perception tasks, two pieces were reused in the experiment in order to assess intra-rater reliability.  

After an analysis of participant performance across the 22 piano works, it was discovered that one of the two pieces used to assess intra-rater reliability resulted in the worst performance of all of the 22 piano works. Therefore, the calculated intra-rater reliability is likely to be skewed or non-representative of participant performance. For
Participants were asked three questions about each pair of musical themes. They were asked to indicate which theme was the first theme and which theme was the second theme in the musical work. Additionally, they were requested to rate how confident they were in their decision on a continuous scale of 1 to 10, using a slider. Before the task started, participants were given the following instructions:

“In this section you will see (or hear) pairs of themes that come from musical scores (or recordings). For each pair, you will be asked to determine which theme you think comes first in the music and which theme you think comes second in the music. In each pair, there will always be one theme that comes first and one theme that comes second. All of the themes have a tonal center of C. After you finish the task, you will be asked to list the criteria you use to make your decisions. Please pay attention to how you are making your decisions.”

After the participants completed both the score-based and recording-based tasks, listeners were asked to rank-order the criteria they used to make their determinations. In addition to the musical criteria from the corpus study (excluding mode and key, which were fixed), five blanks were provided so that participants could include other criteria in the rankings. They were also asked whether first or second themes were easier to classify (generalizable/had more of a prototype) and whether it was easier to classify first and second themes from the musical score or from sound recordings. A last question was posed where participants were asked to comment on the relationship between thematic

these reasons, it is not reported in the results section. Additionally, the confidence level between the test-retest samples changed very little. Overall, the average confidence level for the original discrimination was 5.74 (out of 10) and the average confidence level for the retest discrimination was 5.83 (out of 10).
order (first and second themes) and thematic hierarchy (main theme versus subservient theme).

Finally, after all of this was completed, participants were asked basic demographic questions about their musical experience – estimated years of music theory training, years of instrumental or vocal training, amount of time per week spent listening to classical music, whether or not they were familiar with sonata-allegro form, and their familiarity with Finale and MIDI recordings.

Results

Demographics

Forty-four participants, primarily from the Ohio State University School of Music, took part in the experiment. The median age was 20 (range from 19 to 39). Of these participants, twenty-one (48%) were female. The participants exhibited a wide range of musical training, with a range of 0-25 years of formal music theory training (mean = 3.66). Six participants had one or fewer years of music theory training. Participants had a range of 1-30 years of formal instrumental or vocal training (mean = 9.43). About half of the participants claimed that they had learned sonata-allegro form (48%) and the majority of the participants had medium or no experience with Finale or MIDI (77% had medium to no Finale experience, 82% had medium to no experience with MIDI).

Statistical Analysis

Between the two parts of the study (score-based and recording-based perception), each subject made 44 judgments about first and second themes, with an additional 4
judgments for test-retest trials. With 44 participants, this means that 1936 judgments about first and second themes were made (2112 including test-retest trials). Of these judgments, participants selected the correct answer 57% of the time (1094/1936), which was significantly different from chance, $X^2(1) = 32.802, p < 0.00001$, as shown in Figure 11. Although this is a small effect, the results are consistent with the hypothesis that, on average, participants are able to use structural information from the music to determine which theme from a pair comes first in a musical work. Previous familiarity with the musical works was ascertained from the confidence level.\textsuperscript{34} There was a slight negative-skew in the distribution of confidence ratings of correct decisions and a slight positive-skew in the distribution of confidence ratings of incorrect decisions. However, a person’s confidence in their decision did not seem to dramatically affect the results. The average confidence level for correct decisions was 5.8 out of 10, whereas the average confidence level for incorrect decisions was 5.4 out of 10.

\textsuperscript{34} If a person was 100% confident in his or her determination of first and second themes, it was assumed that he or she was familiar with that musical work. However, it is possible that confidence ratings represent musical training rather than familiarity with the musical works. Given the wording of the question, there was no way to ascertain which of these features confidence level was truly measuring.
Figure 11. Overall judgments of first and second musical themes. Participants were able to discriminate first and second themes in the absence of the original musical context 57% of the time. Although the effect size is small, the effect was statistically different from chance.

**Score-Based Task**

One of the _a priori_ hypotheses was that participants could distinguish first and second themes by looking at the information in the musical score. The data suggest that participants are able to do this at a level statistically different from chance, with 56% of the judgments (539/968) classified as the correct response, $X^2(1) = 12.50, p = 0.0004$. The results are shown graphically in the right half of Figure 12. The corollary to this hypothesis was that participants should perform better when the provided scores include dynamic and articulative information than when the given scores do not include this information. The individual differences from both participant groups were combined to use McNemar’s test, a nonparametric $X^2$ test used to compare differences in categorical data. The results were consistent with this corollary hypothesis, with 56.6% correct judgments in the scores with dynamics and articulation and 54.8% correct judgments in the scores without dynamic and articulation markings. Although these values are close in
magnitude, they are statistically different using McNemar’s test, $X^2(1) = 5.915, p = 0.015$. The results are graphically depicted in the right half of Figure 12.

![Figure 12](image)

Figure 12. Participants perform slightly better in situations when the music contains articulation and dynamic information (“With D.A.”) than when the music does not contain this information (“Without D.A.”) for both the score-based and recording-based judgments.

**Recording-Based Task**

The other main *a priori* hypothesis was that participants could categorize first and second themes when listening to MIDI sound recordings. As shown in the left half of Figure 12, participants were able to perform this task, with 57% of the judgments being the correct answer (555/968), $X^2(1) = 20.831, p < 0.0001$. The corollary to this hypothesis was that participants would be better at categorizing themes from sound recordings when dynamics and articulation were included than from sound recordings without this information. The results of this corollary hypothesis were significant according to McNemar’s test, with 58.9% correct judgments in the recordings with
dynamics and articulation and 55.8% correct judgments in the recordings without
dynamics and articulation, $X^2(1) = 9.820, p = 0.0017$. The results are shown graphically
in the left half of Figure 12.

Participants performed with a slightly higher accuracy in the recording-based task
(57.3% correct judgments) than in the score-based task (55.7% correct judgments),
McNemar’s $X^2(1) = 16.413, p < 0.0001$. This is graphically depicted in Figure 13. In
response to a question about whether it was easier to classify first and second themes
from the scores or the recorded excerpts, 61% (27/44) of the participants indicated that
the recording-based task was easier than the score-based task.

Results by Expectation Group

Recall that the first and second musical themes were compared in terms of their
dynamics and articulation. The differences between the first and second themes on these
two features were then used to categorize the thematic pair according to the degree to which they followed the expectations generated from the results of the corpus study. Thematic pairs where the first themes were louder and more staccato than the paired second themes followed the trend exhibited by a large body of music and were labeled “Expected.” Contrastingly, thematic pairs where the first themes were quieter and more legato than the paired second themes followed the reverse trend and were consequently deemed “Opposite.” It was anticipated that participants should correctly answer the questions in the “Expected” category but incorrectly answer the questions in the “Opposite” category. However, participants performed better in the “Opposite” category than in the “Expected” category, shown in Figure 14. One possible reason for this is because participants consistently performed badly on theme-judgments in one or two of the works in the Expected category. Another possible reason for this unexpected finding was an error in the encoding of the themes from Beethoven’s Sonata No. 22, Mvt 1. It was discovered after the experiment concluded that the themes of this sonata retained their initial I—V harmonic relationship. In order words, the transposition of the second theme was done incorrectly. As the participants were able to use key heuristics to answer this question, the Beethoven was subsequently removed from the sample. After removing the Beethoven sonata from the sample, there was no indication that participants could differentiate first and second themes in the Opposite category. In this category, participants selected the correct theme 50% of the time, $p = 0.962$. The results are discussed further in Appendix C.
Figure 14. Participant performance across the four categories of thematic stimuli. Participants correctly discriminated first and second themes in the expected category 53.8% of the time, but this failed to reach statistical significance, $X^2(1) = 3.03, p = 0.082$. In the opposite category (with the Beethoven included), participants were correct 56.4% of the time, $X^2(1) = 8.76, p = 0.003$. In the conflicting dynamics category, participants were correct 57.0% of the time, $X^2(1) = 8.74, p = 0.003$. In the conflicting articulation category, participants were correct 59.3% of the time, $X^2(1) = 15.28, p < 0.0001$.

The category of musical stimuli (Expected, Opposite, etc.) was determined only with regard to articulation and dynamic markings: it did not take the other structural criteria into account. Average values for the dynamics, articulation, pace, nPVI, and interval size for first and second themes are shown in Table 7. In all four categories of music (Expected, Opposite, Conflicting Dynamics, and Conflicting Articulation), the pace of the first theme was faster than the pace of the paired second theme, consistent with the expectations from the corpus study. Similarly, in all but the Conflicting Dynamics group, the first theme had a higher nPVI value than the paired second theme. Again, this is consistent with the expectations from the corpus study. First and second themes did not have consistent interval sizes across the four categories. First themes had
greater average interval sizes in the Conflicting Articulation and Opposite groups, but smaller average interval sizes in the Expected and Conflicting Dynamics groups.

The fact that participants performed better in the Opposite category than in the Expected category could show that dynamics and articulation were probably not the only features used to determine if a theme comes first or second in the music. Indeed, it could indicate that pace and rhythmic smoothness are equally important for judgments about first and second themes as are dynamics and articulation.

<table>
<thead>
<tr>
<th></th>
<th>Expected</th>
<th>Conflicting Articulation</th>
<th>Conflicting Dynamics</th>
<th>Opposite</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Dynamic</td>
<td>6</td>
<td>5.6</td>
<td>3.2</td>
<td>2.833</td>
</tr>
<tr>
<td>B Dynamic</td>
<td>2.667</td>
<td>2.8</td>
<td>5.6</td>
<td>6.0</td>
</tr>
<tr>
<td>A Articulation</td>
<td>3.5</td>
<td>2</td>
<td>4</td>
<td>2.167</td>
</tr>
<tr>
<td>B Articulation</td>
<td>2.167</td>
<td>3.6</td>
<td>2.6</td>
<td>4</td>
</tr>
<tr>
<td>B Pace</td>
<td>8.122</td>
<td>7.852</td>
<td>6.232</td>
<td>7.922</td>
</tr>
<tr>
<td>A nPVI</td>
<td>45.65</td>
<td>36.38</td>
<td>35.63</td>
<td>49.39</td>
</tr>
<tr>
<td>B nPVI</td>
<td>27.26</td>
<td>30.61</td>
<td>39.90</td>
<td>27.25</td>
</tr>
<tr>
<td>A Interval</td>
<td>1.878</td>
<td>2.752</td>
<td>2.184</td>
<td>3.207</td>
</tr>
<tr>
<td>B Interval</td>
<td>2.407</td>
<td>1.766</td>
<td>3.06</td>
<td>2.093</td>
</tr>
</tbody>
</table>

Table 7. Summary data for the piano works in the expected, opposite, and two conflicting categories. The values shown for all of features are the mean values. “A” refers to first themes and “B” refers to second themes.

Listener Attributes

A few post hoc tests were conducted regarding characteristics of the listeners. The first consideration regarded the amount of music theoretical training in which the participants had engaged. The mean years of theory training in the participant group was 3.66 years. It was thought that music theory graduate students and upper level undergraduate music theory majors may perform better on the tasks. The participants
were accordingly split into two groups. The “Less Theory” group (n = 34) consisted of those who had less than 4 years of music theory training, whereas the “More Theory” group (n = 10) consisted of those who had 4 or more years of training. As expected, the “More Theory” group performed better on the tasks (59.5% correct) than did the “Less Theory” group (55.6% correct), McNemar's $X^2(1) = 173.65$, Bonferroni-corrected $p < 0.001$. The participants with higher musical training performed better on the recording task (61.8% correct) than on the score-based task (57.3% correct), Bonferroni-corrected $p < 0.05$. The participants with less training also performed better on the recording-based task (56.0% correct) than on the score-based task (55.2% correct), although the performances were closer in magnitude, Bonferroni-corrected $p < 0.05$.

The second post hoc test examined whether or not participants with training in sonata theory (n = 23) would do better on the task than participants without training in sonata theory (n = 21). Those with sonata form training performed with a somewhat higher accuracy on the tasks (58.7% correct) than did those without sonata form training (54.5% correct), although this difference was not statistically significant, Bonferroni-corrected $p = 0.126$.

It could be assumed that listening to and engaging with music on a regular basis probably also contributes to perceptual judgments about first and second themes. Accordingly, two more tests were carried out. Participants were divided by the amount of instrumental training they had received. The average amount of training in the participant groups was 9.4 years of instrumental instruction. The participants were therefore split into a “High Training” and “Low Training” category, with those receiving 10 or more years of training in the high condition (n = 24) and those receiving less than 10 years of
training in the low condition (n = 20). The groups performed equally well on the tasks (mean of Low Training = 55.7% accuracy, mean of High Training = 57.5% accuracy, Bonferroni-corrected \( p = 2.83 \)).

Finally, participants were split into two categories based on how often they spent listening to classical music per week. A mean split was performed (mean = 4.33 hours), resulting in “High Listening” (n = 31) and “Low Listening” (n = 13) categories. Those in the high listening category performed significantly better on the task (59.4% correct) than did those in the low listening category (55.3% correct), McNemar’s \( \chi^2(1) = 76.169 \), Bonferroni-corrected \( p < 0.00001 \).

Of course, we must be cautious interpreting these results. There is a certain arbitrariness in dividing a continuous variable into categories, even when well-intentioned. Moreover, many of the groups performed similarly to each other, despite the statistical significance between the groups. Two important trends do emerge, however. Those with more theory training and those who regularly listen to classical music performed the best on the discrimination tasks.\(^{35}\) This follows basic intuition about musical training.

Criteria Used in Decisions

After the subjects finished the two tasks, they were asked to rank order the criteria they used to make their determinations. In addition to the criteria from the corpus study, five blanks were provided so participants could list their own criteria. The results are shown in Table 8. Of the five features hypothesized in the corpus study, rhythmic

\(^{35}\) The correlation between these two groups was -0.09, which means that they likely refer to two different sets of participants.
smoothness (an indication of nPVI) was considered to be the most important feature to aid in discrimination between first and second themes. This is interesting, as it is the only hypothesis that failed to reach significance in the corpus study. However, it could be that the participants understood “rhythmic smoothness” as meaning something different than the nPVI itself. Articulation and dynamics were consistently rated on the lower half of the list of criteria. This is not surprising, as only half of the samples seen by participants contained dynamic and articulation information (the “With DA” condition, as opposed to the “Without DA” condition).

The category marked “Other” was also considered to be important in making the determinations of first and second themes. Participants listed features such as the note density (texture), harmonic complexity, melodic range, cadence type, phrase length, type of meter (simple or compound), speed, and difficulty (virtuosity) as features that helped make the determinations between first and second themes. These additional features are listed in Table 9.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Scores</th>
<th>Recordings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articulation</td>
<td>3.56</td>
<td>3.80</td>
</tr>
<tr>
<td>Dynamics</td>
<td>3.47</td>
<td>3.13</td>
</tr>
<tr>
<td>Pace Value</td>
<td>2.68</td>
<td>2.95</td>
</tr>
<tr>
<td>Interval Size</td>
<td>4.15</td>
<td>4.12</td>
</tr>
<tr>
<td>nPVI</td>
<td>2.22</td>
<td>2.52</td>
</tr>
<tr>
<td>Other (note density/texture, difficulty, harmonic complexity, melodic range)</td>
<td>2.77</td>
<td>2.23 (note density, speed, type of meter, harmonic complexity)</td>
</tr>
</tbody>
</table>

Table 8. Rank-ordered criteria used to determine which theme from a pair comes first or second in the music. The mean importance of each criterion is shown, where the smallest number means it is considered to be the most important (closest to rank number 1).
Table 9. Additional criteria listed by participants as features that helped them perform the discrimination task. Criteria with no mentions in each task are italicized.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Recordings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variations</td>
<td>Variations</td>
</tr>
<tr>
<td>Note Density/Texture</td>
<td>Note Density/Texture</td>
</tr>
<tr>
<td>Sequences</td>
<td>Sequences</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Elaboration</td>
</tr>
<tr>
<td>Declamatory Style</td>
<td>Declamatory Style</td>
</tr>
<tr>
<td>Harmonic Complexity/Stability</td>
<td>Harmonic Complexity/Stability</td>
</tr>
<tr>
<td>Difficulty vs Simplicity</td>
<td>Difficulty vs Simplicity</td>
</tr>
<tr>
<td>Other musical markings (eg Pedal)</td>
<td>Other musical markings (eg Pedal)</td>
</tr>
<tr>
<td>Cadences</td>
<td>Cadences</td>
</tr>
<tr>
<td>Modulation</td>
<td>Modulation</td>
</tr>
<tr>
<td>Phrase Motion/Harmonic Rhythm</td>
<td>Phrase Motion/Harmonic Rhythm</td>
</tr>
<tr>
<td>Melodic Range</td>
<td>Melodic Range</td>
</tr>
<tr>
<td>Chromaticism/Accidentals</td>
<td>Chromaticism/Accidentals</td>
</tr>
<tr>
<td>Mode</td>
<td>Mode</td>
</tr>
<tr>
<td>Melodic Smoothness</td>
<td>Melodic Smoothness</td>
</tr>
<tr>
<td>Speed/Tempo/Time of Listening</td>
<td>Speed/Tempo/Time of Listening</td>
</tr>
<tr>
<td>Meter Difference (simple/compound)</td>
<td>Meter Difference (simple/compound)</td>
</tr>
<tr>
<td>Phrase Length</td>
<td>Phrase Length</td>
</tr>
<tr>
<td>Phrase Completeness</td>
<td>Phrase Completeness</td>
</tr>
</tbody>
</table>

To gain further insight into the importance of individual features in discriminating first and second themes, a post hoc regression analysis was conducted. The aim was to predict judgment type (correct or incorrect) from the measured features in both the first and second theme types. When put in a regression analysis, none of the features reached significance. This relays an important point about the perception of any musical structure. In a regression analysis, each feature’s influence is considered independently. A regression weight indicates the influence of a variable after its shared variance (with other variables) is removed from the equation. In the case of themes, a regression weight for “A Dynamics” would mean that the dynamics for the first theme affect perceptual judgments, given that the first and second themes are equal in their articulation, pace
value, average interval size, and nPVI. The fact that none of the regression weights were significant does not indicate that they do not contribute to judgments about which themes should be considered first or second. Instead, it means that no feature in isolation can be used to predict first and second themes outside of the musical context. Musical features must be taken in combination in order to make these judgments. The reader is encouraged to turn to Appendix C, where specific musical features of first and second themes in four musical works are compared with overall performance on the discrimination task.

Importance of Harmony

A modified version of the perceptual task was conducted after the data analysis from the original perceptual study was carried out. Recall that when participants were asked to describe the criteria they used to make their decisions, many participants indicated that they used harmonic considerations. To investigate the importance of harmony on judgments about first and second themes, the perceptual task was modified. In this modified task, the same sample of musical works was utilized. However, the harmony was removed so that the musical excerpts consisted of only a single melodic line (see Figure 15). The protocol of the task remained the same in this new version.

Figure 15. First theme (top line) and second theme (bottom line) from Dvořák’s, Op. 54, No 3. The first theme with its accompaniment is shown in Figure 10.
Twelve participants completed this modified task, none of whom had taken the original perceptual task. The average age was 20 years (range from 18-25), and 58% were female (7/12 participants). The participants had an average of 4.63 years of music theory training (range from 1.5-14 years) and an average of 8.5 years of instrumental or vocal training (range from 1-14). The results are not consistent with the hypothesis that participants can differentiate first and second themes, with 48% accuracy in judgments, $p = 0.408$. There was no difference in performance for music with or without dynamics and articulation, $X^2(1) = 0.329, p = 0.566$. Participants performed equally well on the score-based and recording-based tasks, $X^2(1) = 0.291, p = 0.589$. A median split was performed to divide participants into groups with more and less music theory training. Both of these groups performed equally well on the tasks, $X^2(1) = 0.307, p = 0.580$.

The results seem to indicate that participants are not able to distinguish first and second musical themes without access to the underlying harmony. However, an interesting trend arises when the music was split by its category (Expected, Conflicting Articulation, Conflicting Dynamics, Opposite). For music in the Expected group, participants performed the task with a 63% accuracy, $X^2(1) = 4.52, p = 0.03$. For music in the Opposite and Conflicting Articulation categories, participants did not perform better than chance (Opposite: $X^2(1) = 0.645, p = 0.422$; Conflicting Articulation: $X^2(1) = 0.078, p = 0.779$). Participants performed significantly worse than chance in the Conflicting Dynamics category, $X^2(1) = 4.65, p = 0.031$.

Although more research needs to be conducted to make any conclusions, one possibility bears mentioning. The results are consistent with the interpretation that participants are able to use surface level features to differentiate first and second themes.
in cases when the music conformed to previous expectations of how dynamics and articulation should differ. However, when articulation and dynamics did not conform to previous expectations, participants could not complete the task without access to the underlying harmonies. This hypothesis must be directly tested in future research and provides exciting avenues for research on musical themes.

Discussion

The goal of the current study was to determine whether people could differentiate between a pair of musical themes using perceptual factors. Overall, participants were able to use structural musical features to differentiate first and second themes from the same musical work; they could do this both from sound recordings and from musical scores. In general, participants performed better in the conditions where dynamics and articulation were available than in conditions without these features; however, the performance across these two conditions was similar. In general, performance was similar across many of the conditions of the study. For example, in both the score-based task and recording-based task, accuracy in judgments was consistently around 57-60%. This is significantly different from chance, but is small in magnitude, indicating that the task was difficult. Participants with more theory training and those who regularly listen to classical music performed the best on the discrimination tasks. Because of the difficulty of the task, future research should primarily recruit participants with a high level of music training. Future research should also utilize a larger and more representative sample of music in order to investigate which factors may make the discrimination task easier or more difficult.
Score-reading and close-listening are related, but distinct musical skills. Despite
the nearly identical performance on the score-based and recording-based tasks,
participants seem to indicate that the recording-based task was easier than was the score-
based task (61% indicated that the recording-based task was easier, but only 18%
indicated that the score-based task was easier). Future research should be conducted to
ascertain if this difference in comfort between score-based and recording-based tasks is
consistent across a wider body of participants or is sample specific.

In this perceptual study, we have seen that participants are able to use structural
features of the music to differentiate between first and second musical themes. In the
final chapter, we will summarize the results of the two major studies of the thesis. Further
observations and discussion will also be presented.
Chapter 5: Conclusions

The aim of this thesis was to examine specific properties of musical themes in Western art music. Two methods of examining contrasting expressive content within first and second musical themes were presented. Although contrasts between themes are typically investigated using standard musical analysis, the current study used an empirical-based methodology so that common claims about themes could be assessed empirically.

Chapter 2 explained that theorists have noted differences between first and second musical themes over the past 300 years. First themes have been characterized using descriptive terms like masculine, strong, and energetic. They have also been associated with structural features such as loud dynamics, fast tempo, staccato or detached articulations, and disjunct rhythms. Common descriptions of second themes are terms such as feminine, lyrical, and gentle. Second themes are thought to contain features like quiet dynamics, slow tempo, legato articulations, and smooth rhythms. Contrasts between first and second themes are often taught in undergraduate music theory classes and have become standard features in many music theory textbooks.

The goal of Chapter 3 was to empirically test these extant generalizations in a large body of music and to examine how treatment of these two theme types differed over stylistic musical periods. The chapter described a corpus study, or the study of a large
body of music, which examined structural features of classical music themes over a period of 300 years. The music examined consisted of a random sample of 1063 pairs of first and second themes taken from Barlow and Morgenstern’s *Dictionary of Musical Themes*. The Humdrum Toolkit was used to measure the encoded musical themes on four features of interest: mode, average interval size, pace value, and rhythmic smoothness. Additionally, articulation and dynamic levels were manually encoded. Nonparametric tests were used to compare the construction of first themes with the construction of their paired second themes. In our sample, compared to first themes, second themes were significantly less likely to be in the minor mode, but were more likely to be legato, make use of a slower tempo, use smaller intervals, and to involve a quieter dynamic level. These differences were similar in musical works spanning the seventeenth through the twentieth centuries.

The aim of Chapter 4 was to assess whether musically trained listeners could perceive structural differences in musical excerpts. The goal was to determine whether musicians could differentiate between a pair of musical themes using features of the musical construction. The stimuli consisted of first and second themes from 22 works for solo piano. The study was conducted in two stages: in one stage, participants used musical scores, and in the other stage, they listened to sound recordings. In the experiment, participants were presented with two musical themes. They were asked to indicate which theme they believed came first in the music and to rate the confidence of their response on a 10-point continuous scale. In some cases, participants received music with dynamics and articulation included. At other times, participants received music with this information removed. Nonparametric tests were used to make the comparisons. The
results of the perceptual study are consistent with the hypothesis that musicians can
differentiate first and second musical themes using features of the music itself, at an
accuracy rate of about 57%. They are able to do this even when the themes are isolated
from the original musical context. We saw that participants performed slightly better on
the recording-based task than on the score-based task. Participants tended to perform
better when articulation and dynamics were included than when these features were
withheld, although the difference between these conditions was small in magnitude.

Limitations and Future Directions

There are certain limitations of the current study that must be acknowledged. One
of the intentions of the study was to examine a large and representative body of musical
works. In order to do this, an encoded version of the Barlow and Morgenstern Dictionary
of Musical Themes (1948) was used to randomly generate a sample of themes. The use of
the Dictionary prevents a priori preconceptions of the current author from affecting the
musical sample. However, the Dictionary still contains biases of Barlow and
Morgenstern, especially bias towards music written in the nominal Classical and
Romantic eras. Additionally, there was little effort made to independently corroborate or
verify the designations of “First Themes” and “Second Themes” within the Barlow and
Morgenstern Dictionary. It is possible that through close musical analysis, errors in
Barlow and Morgenstern’s labels would be detected. In summary, a limitation of this
thesis is the unknown correspondence between the sample of themes used in this study
and the population of all Western-art musical themes.
Another limitation of the current study is that the research focused on musical works that employed only two themes—works with more than two themes were not considered. It is possible that the relationship between first and second themes will differ when there are also other themes in the work. Additionally, first and second themes were only examined in the context of their first appearance in the musical work; the development of these themes throughout the musical work was not considered.

Furthermore, the results of these studies are limited to the features that we measured in the corpus: modality, durational pace, average interval size, rhythmic smoothness, articulation, and dynamic markings. We operationalized these measures in order to conduct empirical tests. Other researchers may have operationalized these terms in different ways, which could potentially affect the results of the studies. In addition to these six measured features, there are also other ways that themes might be similar or contrasting. These other features provide rich avenues for future work about musical themes. For example, in the later Romantic period, it was common for some composers to experiment with “bigger” timbres, larger orchestras, and programmatic content. These composers likely used factors such as instrumentation, harmony, and orchestral treatment to shape possible differences (or similarities) between first and second themes. With new musical experimentation came new stylistic differences, a story that fits well with what we already know about musical and historical change. Future studies should take into account these additional features when examining differences in expressive musical thematic content.

The perceptual study also contained limitations. The sample of musical works consisted entirely of piano music written from the late eighteenth century to the early
twentieth century. The results of the study may not generalize to music written for other instruments or from other time periods. The participants in the study were also likely non-representative of most musicians. Most of the participants were undergraduate music majors from a single institution, namely the Ohio State University, where they received similar music theory training. The study should be extended to musicians with a wider range of music theory training and who have received their training at other institutions.

One potential avenue for future research is how the order of musical themes relates to their prominence in the musical work. The themes studied in this thesis are classified by the order in which they appear in the musical work; first themes are presented before second themes. The order in which themes appear may not be related to how prominent (or important) each theme sounds. In some circumstances, the first theme may be considered to be more important; in other cases, the second theme may be considered to be more important. Future research should be conducted in order to determine the relationship between these two constructs (thematic order and prominence). To begin this research, a task could be conducted where participants list terms they associate with thematic prominence. A content analysis could be carried out on the terms named by the participants. From these results, an operationalization of thematic prominence (or importance) could be reached.

One final consideration that deserves further study is how the information from the musical score differs from performance practices. Especially in the case of articulation, there are stylistic conventions that composers would not have needed to write in the score, with the expectation that performers will play the music in a certain way. Certainly, Baroque and Romantic articulation conventions are different and are not
always marked in the score (for example, bowing decisions, and how the chosen speed of a section affects what types of articulation are possible). It is possible that some of the “textbook” generalizations about first and second themes arose from common performance practices, in addition to the notated musical score. Since the current study only used the musical score to characterize first and second themes, future research might also take into account conventional performance practices. An example of how this could be addressed is by modifying future perceptual tasks. The perceptual study described in this thesis utilized Finale-generated MIDI performances. By using real performances of music in a perceptual study, listeners would have access to additional information, which could alter the way they perform the tasks.

The hope of this thesis was to contribute to the centuries-long discussion regarding musical themes. In order to conduct empirical research, it was necessary to rely on assumptions and to make use of methods that have limitations. By using two empirical perspectives – a corpus study and a perceptual study – I hoped to provide complementary perspectives regarding the use of contrast in first and second musical themes. The methods used provide some empirical support for claims made for hundreds of years about contrasting musical themes. Additionally, the use of empirical methods has helped to provide some insight into which musical and acoustic factors may be perceived by musicians in both visual and auditory settings. Furthermore, the current work can be extended to future research projects. There have been few research studies regarding the perception of musical themes. Future studies can build on the preliminary findings presented here to investigate more fully which factors of the musical structure are perceived by musicians.
References


Appendix A: Evaluation of Barlow and Morgenstern’s Themes

The purpose of this Appendix is to understand how Barlow and Morgenstern labeled themes in their *Dictionary of Musical Themes* (1948). First, a table will be provided in order to detail the correspondence between this author’s analysis and Barlow and Morgenstern’s analysis. Next, individual examples of musical themes will be shown graphically. Following that, there will be a short discussion of Liszt’s *Transcendental Etude No. 4*, Holst’s *Mercury, the Winged Messenger*, and Debussy’s *Prelude to the Afternoon of a Faun*. Finally, a graphical depiction of Mozart’s *Piano Concerto in Eb K. 482* will show differences between analyses by Caplin, Hepokoski and Darcy, and Barlow and Morgenstern.

An analysis of a few works in the primary sample of music used in Chapter 3 was conducted. The works analyzed were chosen so that works from different historical periods of music would be examined. The list of works examined is given in Table 10, where it can be seen that the semi-random sample was heavily biased towards piano works and orchestral works. For each work, the correspondence between my analysis and the analysis of Barlow and Morgenstern is listed. Where there are discrepancies between the two analyses, a brief description is provided of how the analyses differed.
Table 10. A comparison of thematic content in musical works. The agreement levels are ranked in terms of how well Barlow and Morgenstern’s analysis corresponds with my own analysis. From high to low agreement, the levels are ordered in the following way: total agreement, high agreement, general agreement, some agreement.

<table>
<thead>
<tr>
<th>Composer</th>
<th>Musical Work</th>
<th>Date</th>
<th>Type</th>
<th>Agreement with Dictionary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bach, J.S.</td>
<td>Italian Concerto, Mvt 3</td>
<td>1735</td>
<td>Keyboard</td>
<td>Total agreement.</td>
</tr>
<tr>
<td>Bach, J.S.</td>
<td>Well Tempered Clavier II, Prelude 3</td>
<td>1738</td>
<td>Keyboard</td>
<td>Some agreement – I would have broken the “themes” into smaller motivic units</td>
</tr>
<tr>
<td>Handel, G.</td>
<td>Concerto Grosso in D Minor, Op. 6 No. 10, Mvt 5</td>
<td>1739</td>
<td>String Orchestra</td>
<td>Some agreement – I would have broken the “themes” into smaller motivic units</td>
</tr>
<tr>
<td>Mozart, W.A.</td>
<td>Concerto in G, K 453, Mvt 1</td>
<td>1784</td>
<td>Piano and Orchestra</td>
<td>General agreement</td>
</tr>
<tr>
<td>Mozart, W.A.</td>
<td>Concerto in Eb, K 482, Mvt 1</td>
<td>1785</td>
<td>Piano and Orchestra</td>
<td>Some agreement – examined in detail below.</td>
</tr>
<tr>
<td>Beethoven, L.</td>
<td>Sonata Op. 2 No. 2, Mvt 1</td>
<td>1794</td>
<td>Piano</td>
<td>Some agreement – the beginning of the themes agree with designations given by Caplin (Caplin, 1998). Caplin believes that there is a “Subordinate Theme 2” after the second theme in the Dictionary (“Subordinate Theme 1”). Hepokoski and Darcy (2006) call this “Subordinate Theme 2” the “Closing Theme.”</td>
</tr>
</tbody>
</table>
Some discrepancies were found between my analysis and Barlow and Morgenstern’s analysis in terms of when the musical themes ended. Barlow and Morgenstern were likely constrained by space issues in their Dictionary, where each theme takes up one line of the Dictionary. Therefore, the length of the themes – as deemed by Barlow and Morgenstern – may have been determined by the amount of music that could fit on one notated line. In Baroque-era music, Barlow and Morgenstern labeled multiple sequences of a small motive as one musical theme, presumably so that...
the theme was stretched to be a certain length. Figures 16 and 17 provide examples of this. For Classical works in sonata form, the themes labeled by Barlow and Morgenstern tended to be significantly shorter than themes generated by a musical analysis. Figure 18 compares an analysis of a Classical era theme by Caplin (1998) and by Barlow and Morgenstern. Additionally, sometimes Caplin, Hepokoski, and Darcy posit the existence of other themes ("Subordinate Theme 2," "Closing Theme") that the Dictionary does not include. An example is shown in Figure 20.

The other discrepancy found in the Dictionary regards works written in large ternary form or works labeled minuet/scherzo and trio. One might expect a priori that the first theme of those works would correspond to the A section or minuet/scherzo and the second theme of those works would align with the B section or trio section. In the Dictionary, this is typically the way Barlow and Morgenstern labeled themes in Romantic-era works, like the second movement of Mahler’s First Symphony and in Chopin’s Nocturne Op. 48, No. 1. However, for Classical-era works written in these forms, it appears that Barlow and Morgenstern only attributed one musical theme to the entire movement. Examples of this include the third movement of Beethoven’s Piano Sonata in E Flat, Op. 7, the third movement from Beethoven’s Violin Sonata Op. 24, and the second movement of Mozart’s Piano Sonata in A, K 331. For the purposes of this thesis, having an inconsistent labeling system for works in these forms results in a somewhat skewed sample. However, the inclusion of over one thousand pieces of music

---

36 Upon close analysis, it seems that Barlow and Morgenstern were also inconsistent in their thematic labeling of rondo form works. Some works in rondo form were labeled with only one theme for the movement, whereas others were given up to four themes, presumably corresponding to some of the couplets. Further analysis should be conducted to understand how Barlow and Morgenstern parsed rondos into musical themes.
in the main study hopefully compensates for such additional biases. More research needs to be conducted about how Barlow and Morgenstern labeled works of different musical forms.³⁷

Figure 16. Bach’s *Well Tempered Clavier Book 2, Prelude 3*. Here we can see that what Barlow and Morgenstern identify as a theme consists of a sequence of a musical motive (first theme) and a subject and countersubject (second theme). I suggest that the inclusion of the extra musical material is so that the music fills an entire line of the *Dictionary*.

³⁷ For works written in sonata form, it appears that Barlow and Morgenstern labeled “first themes” as the traditional primary theme and “second themes” as the traditional secondary theme (the theme in the subsidiary key). This author found no indication that “second themes” ever corresponded to transitions or to a second primary theme within a primary thematic group.
My analysis, first motive
Barlow and Morgenstern, First Theme

My analysis, second motive
Barlow and Morgenstern, Second Theme

Figure 17. Handel’s *Concerto Grosso in D Minor, No. 10, Mvt 5*. Here we can see that what Barlow and Morgenstern identify as a theme consists of a sequence of a musical motive. I suggest that the inclusion of the extra musical material is so that the music fills an entire line of the *Dictionary*.
Beethoven’s Piano Sonata Op. 2 No. 2, Mvt 1, Theme 1. Here we can see that what Barlow and Morgenstern identify as the First Theme corresponds to the antecedent of Caplin’s Main Theme. Similar trends of shortening musical themes can be seen across many Classical-era works in the Barlow and Morgenstern Dictionary.

Analysis of Liszt, Holst, and Debussy

Three examples provide additional insight into how Barlow and Morgenstern may have labeled musical themes. The first example is the Transcendental Etude No. 4 (“Mazeppa”) by Franz Liszt. In this musical work, Barlow and Morgenstern identify two musical themes, shown in Figure 19. The melodies of the first and second theme are almost identical. However, the accompaniment that surrounds the melodies is wildly different, as is the overall tempo of the two designated themes. When a listener hears these two themes in the musical work, he or she will likely notice that there is a strong...
affective difference between the two musical themes. However, it is possible that a musician will notice the relationship between the so-called First Theme and Second Theme. In my own analysis, I would indicate that these two themes are related, with the “Second Theme” perhaps being labeled as a variant of the First Theme. The fact that Barlow and Morgenstern differentiate between these two themes could indicate that, to them, sensitivity to musical characteristics like accompaniment, tempo, and register are equally important as are considerations about the musical notes.

Figure 19. Liszt’s *Transcendental Etude No. 4 “Mazeppa.”*
An analysis of Gustav Holst’s “Mercury, The Winged Messenger,” from *The Planets* was also conducted. “Mercury” is written in rondo form, with clear boundaries between sections. The work starts with an A section, which itself can be divided into two parts. The work opens with Part 1 of this A section, which I deemed the “Whirling Introduction.” In this section, there are a series of memorable and short patterns – clear textures, timbres, contour patterns, and rhythmic patterns – that return later in the work. It is not “thematic” in a melodic sense, but in all other aspects it could act like a musical theme. Barlow and Morgenstern do not label any part of this “Whirling Introduction” as a theme. Instead, their First Theme begins 47 measures into the work, and consists of a clear melodic pattern. This also corresponds with the beginning of Part 2 of the A section. Barlow and Morgenstern’s Second Theme corresponds with the theme at the beginning of the B Section, a designation with which I agree. Therefore, it is possible that Barlow and Morgenstern require a certain type of melodic component for a musical excerpt to be considered as a theme in their *Dictionary*. This would seem to serve the original purpose of the *Dictionary*, which aimed to produce an index of memorable phrases and melodies to be used as a reference tool.

The work *Prelude to the Afternoon of a Faun*, by Claude Debussy, can be used to identify one last aspect of thematic choices in the Barlow and Morgenstern *Dictionary*. The choices for First and Second Themes by Barlow and Morgenstern are ones with which I completely agree. However, within this music, there are additional distinctive melodies played by solo instruments, such as the oboe. These melodies do not come back in the work, but they are memorable, distinctive, and short. Barlow and Morgenstern ignore these melodies and do not designate them as themes. Therefore, it seems that they
do not label melodies or “tunes” as themes; a musical theme may be required to repeat or be developed throughout the musical work.

By examining the Barlow and Morgenstern themes through my own analytical eyes, as well as those of William Caplin, James Hepokoski, and Warren Darcy, we have allowed the musical compositions to speak for themselves. An overall impression of the Barlow and Morgenstern *Dictionary of Musical Themes* indicates that they chose musical themes wisely and in ways that correspond with traditional musical analysis. Of course, a more in depth analysis of works must be undertaken to provide any conclusive evidence about their methodology for labeling musical themes. However, the small-scale analysis conducted here provides hope that the musical themes used in this thesis are sensible choices and are musically meaningful. With this encouraging sentiment, we can move forward with analyzing a large body of musical themes taken from the Barlow and Morgenstern *Dictionary of Musical Themes*.

Analysis of a Mozart Piano Concerto

Mozart’s *Piano Concerto in Eb K. 482, Mvt 1* will be used as a musical example to compare the approaches of Barlow and Morgenstern, Hepokoski and Darcy, and Caplin. The reader should refer to the printed score below. Boxes around the music designate the beginning and end of themes. Hepokoski and Darcy’s analyses are shown in green, Caplin’s analysis is shown in red, and Barlow and Morgenstern’s analyses are shown in purple.

In general, all three sets of scholars agree where themes begin. In the Barlow and Morgenstern *Dictionary*, themes usually last for only a few measures. Hepokoski, Darcy and Caplin all indicate that themes last for longer than a few measures. Additionally, it
can be seen that Barlow and Morgenstern do not consider the transition to be its own theme. Instead, their second theme corresponds with the secondary (subsidiary) themes of Caplin, Hepokoski, and Darcy.

One interesting feature of the Mozart Eb Concerto is that the opening ritornello and the opening solo section contain different first themes. In other words, the first theme of the piano exposition is a different theme from the first orchestral theme. Barlow and Morgenstern’s Dictionary only accounts for the themes present in the ritornello exposition. It ignores the introduction of new material in the opening solo exposition.

Caplin, Hepokoski, and Darcy sometimes label themes differently, as seen in the annotated score. The two sets of scholars have different criteria for determining where themes begin and end, as detailed in their books (Classical Form and Elements of Sonata Theory). These differences are beyond the scope of this thesis.
Figure 20. Mozart’s *Piano Concerto in Eb K. 482, Mvt 1*
Figure 20 continued

Continued
Figure 20 continued
Figure 20 continued
Figure 20 continued
Figure 20 continued
Figure 20 continued
Figure 20 continued
Figure 20 continued
Figure 20 continued

Caplin, Subordinate Theme 1, Part 2 (continues)

Continued
Figure 20 continued
Figure 20 continued
Appendix B: Additional Materials from the Perceptual Study

<table>
<thead>
<tr>
<th>First theme louder, more staccato (“Expected”)</th>
<th>Dvorak, Antonin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waltzes, Op 54, No 3</td>
<td></td>
</tr>
<tr>
<td>Sonata in D K 311, Mvt 1</td>
<td>Mozart, Wolfgang Amadeus</td>
</tr>
<tr>
<td>Movements Perpetuels, No 3</td>
<td>Poulenc, Francis</td>
</tr>
<tr>
<td>Sonata No 6 Op 82, Mvt 1</td>
<td>Prokofieff, Serge</td>
</tr>
<tr>
<td>Rhapsody in E Flat Op 119 No 4</td>
<td>Brahms, Johannes</td>
</tr>
<tr>
<td>Minuet in G</td>
<td>Beethoven, Ludwig Van</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>First theme quieter, more legato (“Opposite”)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sonata No 22 in F Op 54, Mvt 1</td>
<td>Beethoven, Ludwig Van</td>
</tr>
<tr>
<td>Sonata No 23 in F Minor Op 57</td>
<td>Beethoven, Ludwig Van</td>
</tr>
<tr>
<td>&quot;Appassionata,&quot; Mvt 3</td>
<td>Liszt, Franz</td>
</tr>
<tr>
<td>Valse Oubliee</td>
<td>Mendelssohn, Felix</td>
</tr>
<tr>
<td>Songs Without Words, No 23 in A, Op 53</td>
<td>Falla, Manuel de</td>
</tr>
<tr>
<td>No. 5 &quot;Folk Songs&quot;</td>
<td>Chopin, Frederic</td>
</tr>
<tr>
<td>4 Pieces Espagnoles, Montanese</td>
<td></td>
</tr>
<tr>
<td>Scherzo in C Sharp Minor Op 39</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>First theme louder, more legato (“Conflicting Articulation”)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sonata in F, K 332, Mov. 3</td>
<td>Mozart, Wolfgang Amadeus</td>
</tr>
<tr>
<td>La Vie Parisienne, Act II Entr'acte – Piano</td>
<td>Offenbach</td>
</tr>
<tr>
<td>Transcription</td>
<td></td>
</tr>
<tr>
<td>Sonata No 2, Mvt 2</td>
<td>Hindemith, Paul</td>
</tr>
<tr>
<td>Hungarian Rhapsody No 4 in E Flat</td>
<td>Liszt, Franz</td>
</tr>
<tr>
<td>Reverie</td>
<td>Chopin, Claude</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>First theme quieter, more staccato (“Conflicting Dynamics”)</th>
<th>Paderewski</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cracovienne Fantastique</td>
<td></td>
</tr>
<tr>
<td>Sonata No 14 in C Sharp Minor, Op 27 No 2,</td>
<td>Beethoven, Ludwig Van</td>
</tr>
<tr>
<td>&quot;Moonlight&quot;, Mvt 2</td>
<td></td>
</tr>
<tr>
<td>Sonata No 21 in C Op 53 &quot;Waldstein&quot;, Mvt 1</td>
<td>Beethoven, Ludwig Van</td>
</tr>
<tr>
<td>Sonata in B Flat Op 47 No 2, Mvt 3</td>
<td>Clementi, Muzio</td>
</tr>
<tr>
<td>Polonaises Op 40 No 1 in A</td>
<td>Chopin, Frederic</td>
</tr>
</tbody>
</table>

Table 11. A List of Musical Works Used in the Perceptual Study.
Thematic Order and Thematic Hierarchy

In the perceptual study, participants were asked to comment on the relationship between thematic order (first versus second themes) and thematic hierarchy (main theme versus subservient theme). Participants were asked the following question: “Do you think there is a relationship between the order of themes (first theme and second theme) and how important the themes are? Are first themes always the main theme or are second themes sometimes the main theme?” Participant responses are recorded below.

- I think most of the time the first themes are the main theme but there are sometimes exceptions to that where the second theme could be the main theme.
- I think the first theme usually tends to be the main theme.
- I think usually the first theme is the main theme, since it usually reappears in altered forms or exactly as written.
- Usually, the second theme is somehow related to the first.
- I do not know. I’m not familiar with this style of music.
- There is definitely a relationship between the order of themes because there could be a small embellishment of the first theme in the second theme which could easily be determined. The first themes are usually the main themes because they are stated first and usually repeated.
- There is a relationship between themes and no, the first theme is not always the main theme.
- First themes are always the main theme.
- I think first themes are almost always the main theme.
- As far as the piece as a whole goes, theme relation should be something considered and articulated throughout the piece. When isolating the first theme and second theme, there is little justification to what is more appropriate.
- Yes, I do believe that there is a relationship between the order of the themes. If the score looked busier (as in more notes and complicated rhythms), I was more likely to assume it was the second theme. In the listening section, I felt that the tempo played a more important factor than it did in the scores.
- Themes are important by finding the section and form of a piece. The first theme may sometime tell what the second theme (importance) of the piece.
- Not always, typically first themes are simpler in nature and are more susceptible to variation, so on and so forth. Aurally it is easier to pick up on things, tempo, dynamics, so on... also typically in this style, often times structure of whole pieces are fast, slow, fast in structure. Also I recognized a few pieces of Mozart and the reverie piece and a Beethoven piece once I heard them played.
- It depends on the dynamic and how they lead into each other.
• No, either theme can be more important, it just depends on how the piece is laid out.
• Sometimes second themes can be the main themes. It depends on who composed the piece, and when.
• I think the first theme is usually stronger. I think about how the composer needs the listener to be able to recognize it later in the piece. It doesn't have to be louder or have more notes. It needs to be identifiable, which usually means it has a simple yet clear structure without a complex texture. The second theme can usually be less defined. It can even be more of a textural gesture than a melodic theme. I think themes are essential to the structure of music.
• There is no real importance, in fact a second theme can be the main theme of a piece too. It depends on each individual piece.
• I think first themes are the main theme.
• First themes are usually the main theme, but that doesn't mean that they are more important.
• First themes are a bit more memorable, so they're the main theme.
• I think the first theme is almost always the main theme.
• I think both are of equal importance -- it's just that each theme is included to serve a different purpose.
• I suppose first themes tend to be easier to remember, so from a listener's perspective, they tend to be more prominent and important.
• I think the first themes are always the main theme. And second themes should be more "fancy" than first themes.
• I think first themes are usually the main theme, because they tend to contain an idea/ideas which are then repeated in various ways throughout the piece.
• The first theme is typically gonna be simpler. Also the second theme is going to employ aspects of the first theme.
• I perceive there to be a relationship between the order of themes, therefore I look for relationships where they may not necessarily be one. I would say first themes are the main theme most of the time (perhaps 80%) but second themes are the main theme enough times to actually matter.
• I believe that the second theme can sometimes play off of the first theme using different theory and compositional techniques and have a relating (similar or contrasting) idea.
• First or second themes can be the main theme. Or the two can interact, there doesn't have to be a dominant theme. But the relationship between the order of themes definitely does play a role for a listener's ear, because the second theme is usually always related to the first in a way. The first merely establishes something, and the second plays with it, emphasizes it, de-emphasizes it, or any number of other things.
• I believe 'First Themes' play the role of establishing a simpler melodic idea, as compared to the second, which may be more embellished.
• It does not seem like there is a relationship between the order of themes. I think for the most part the first theme is the main theme.
• I think there is a slight relationship in the order of themes. First themes can sometimes be just an intro and the second theme could actually be the main theme of the song.
• The first themes are usually the main theme.
• I feel like the first themes are more often the most important, but can sometimes develop into a more important second theme.
• I think it depends on the type of piece. First themes are most important if they are being carried on and developed throughout the piece, however that is not always the case.
• It seems to be that the first theme is generally more stable than the second.
• I think that the first theme is generally the main theme, though the second one is occasionally as or more important.
• I think there is a relationship, only for the circular reasoning that people are taught that in music school. Because of that, you go into the experiment with the knowledge that first themes SHOULD be louder, more bold, masculine, and major whereas second themes tend to be more frilly, feminine, quieter, and more legato. And in terms of 'main-ness', its normally seen as an interplay between musical ideas as opposed to importance of one over the other. Second themes in their contrasting nature bring out whatever opposite characteristics in the first theme (relatively speaking) just because they might be different.
• I tend to think of the first theme as 'catchier' and the second theme as 'lyrical' but this isn't always the case. I would have thought that I liked first themes better, but you didn't ask me which one I liked more. That might be interesting to look at in the future. In answer to your real question, I don't know.
• I think the answer might depend on the genre and style of music under consideration. If we focus on tonal music (especially in the Classical period), the first theme will be probably more important than (or at least equally important as) the second theme.
The two themes on this page come from the same piece of music.

One of the themes comes first and the other theme comes second.

This theme comes FIRST.

This theme comes SECOND.

This theme comes FIRST.

This theme comes SECOND.

How confident are you about which theme is FIRST?

0 means "not confident at all" and 10 means "I am completely certain"

Confidence: 7.746

Figure 21. Score-based example with accompaniment (no dynamics/articulation)
The two themes on this page come from the same piece of music.

One of the themes comes first and the other theme comes second.

This theme comes FIRST.

This theme comes SECOND.

This theme comes FIRST.

This theme comes SECOND.

How confident are you about which theme is FIRST?

0 means “not confident at all” and 10 means “I am completely certain”

0  1  2  3  4  5  6  7  8  9  10

Confidence

Figure 22. Score-based example without accompaniment (with dynamics/articulation)
Think back to how you performed the two parts of the study:

**How were you able to tell which theme came first?**

Please rank order how important each of these criteria were, when they were available, in deciding which theme came first from the *Musical Score*.

Rank “1” should be considered the most important criterion. If you did not use a criterion listed, do not include it in the rankings.

There are blanks so you can write in any criteria not listed and include it in the rankings.

- Articulation
- Average Interval Size
- Average Note Length/Duration
- Dynamics
- Rhythmic Smoothness
- Other Criterion #1
- Other Criterion #2
- Other Criterion #3
- Other Criterion #4
- Other Criterion #5

Figure 23. Criteria options for the score-based task. An identical question was asked with regard to the musical recordings.
Appendix C: Participant Scores and Quantitative Indices of Musical Works

Of the twenty-two piano works used in the perceptual study, Beethoven’s Sonata No. 22 and Brahms’ Rhapsody in E Flat resulted in the most accurate judgments about first and second themes. The worst performances were in response to a waltz by Dvorak and a Prokofiev sonata. Detailed information about these works can be seen in Table 12. The similar confidence ratings across the works suggest that the participants were relatively equal in their familiarity with the works (around 6 out of 10).

Recall that in Chapter 4, musical stimuli were grouped into four categories, depending on the dynamics and articulation of the first and second themes: Expected, Opposite, Conflicting Dynamics, and Conflicting Articulation. It was expected that participants should perform well on pieces in the Expected category and badly on pieces in the Opposite category. However, we noted that the opposite trend seemed to occur. The two worst performances (in response to the Dvorak waltz and the Prokofiev sonata), were both in the “Expected” category of articulation and dynamics. However, the best performance was in response to a Beethoven sonata, which was in the “Opposite” category of articulation and dynamics.38 This could indicate that participants may not

---

38 As mentioned in Chapter 4, it was discovered after the experiment that the second theme of the Beethoven had been transposed to GM instead of to CM. Therefore, the original key relationships (I—V) were maintained in the perceptual study. In this piece,
have used articulation and dynamics to help them with their discrimination. Therefore, the terms “Expected” and “Opposite” are likely misnomers; whether or not a piece belongs to the “Expected” group does not accurately reflect its musical structure. To examine this idea further, the other measured structural features between these works were compared.

Among these four works, there is a range in musical features, such as dynamics, articulation, pace, nPVI, and average interval size. In the Beethoven and Brahms, the nPVI of the first theme is much higher than that of the second theme. This is not the case in the Dvorak and the Prokofiev, however. Consistent with participant’s rankings of criteria in Chapter 4, the rhythmic smoothness of the two themes may be one of the biggest indicators of which theme comes first in the music. Interestingly, in both the Beethoven and the Brahms, the pace of the second theme was faster than that of the first theme, an opposite trend of what was expected from the results of the corpus study.

Turning to thematic features not discussed by the a priori hypothesis, we can briefly speculate as to what listeners may have been attending when making their discriminations. The musical scores to these four thematic pairs are shown in Figures 24-27. In the Beethoven pair, the first theme clearly prolongs the tonic, which helps solidify the function of main themes, according to Caplin (1998). Therefore, despite the quiet, more lyrical nature of the theme, the tonic prolongation may have helped discriminate the first theme from the more sequential second theme. The Brahms thematic pair had the expected relationship between first and second themes in all attributes except the pace value and therefore the correct overall average is not surprising. Additionally, after the high accuracy rate in judgments about first and second themes could be attributed to this coding error.
transposition of the two themes to CM, both of the themes in the Brahms are strongly rooted in CM. Similar to the Beethoven example, the Dvorak had one theme that was more firmly rooted in “tonic” (recall that both themes were transposed to C in the experiment) than the other. In this case, it is the second theme. Therefore, similarly to the Beethoven example, the tonic prolongation could have been a main feature of interest to the listeners, overriding considerations of dynamics, articulation, pace, interval size, or nPVI. Additionally, the alternating dynamics in the first theme could have obscured the relationship between the primary dynamics of the first theme (forte) and the second theme (pianissimo). The Prokofiev once again follows a similar trend, where the second theme is more clearly centered around a home key, although it is beginning to change harmonic focus by the end of the theme. The first theme in the Prokofiev may have been confusing to some participants, especially ones with less theory training, as it is very dissonant and contains dramatic tritone leaps in the bass. Therefore, overall, a clear establishment or prolongation of a home key seems to be a feature that helps participants in their judgments about first and second themes.\footnote{This is a feature of central importance to thematic construction, according to Caplin (1998).}

A modified version of the perceptual task was conducted, where the harmony from the musical samples was removed. In general, participants were not able to distinguish between first and second themes without the harmony. In terms of the Brahms, Beethoven, Dvorak, and Prokofiev, the accuracy in judgments about first and second themes became more equal, although participants still performed better on the Brahms and Beethoven than they did on the Prokofiev and Dvorak. Table 13 shows the accuracy scores of the modified perception task. The removal of the harmony resulted in
better performances in judgments about the Dvorak and the Prokofiev works – the two
pieces that participants had performed the worst on in the original experiment. However,
the removal of the harmony resulted in worse performances in judgments about the
Brahms and the Beethoven works – the two pieces that participants had performed the
best on in the original experiment. This shows how important harmonic considerations
are likely to be in judgments about first and second musical themes.
Table 12. Descriptions of four of the piano works in the perceptual tasks. People performed the best on Beethoven’s Sonata No 22 in F, Op 54, Mvt 1 and on Brahms’ Rhapsody in E Flat Op 119. In contrast, people performed the worst on Dvorak’s Waltzes Op 54, No 3 and Prokofiev’s Sonata No 6 Op 82, Mvt 1.

<table>
<thead>
<tr>
<th>Musical Work</th>
<th>Overall Accuracy</th>
<th>Accuracy in Recordings</th>
<th>Accuracy in Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dvorak</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Prokofiev</td>
<td>58%</td>
<td>100%</td>
<td>20%</td>
</tr>
<tr>
<td>Brahms</td>
<td>75%</td>
<td>67%</td>
<td>83%</td>
</tr>
<tr>
<td>Beethoven</td>
<td>67%</td>
<td>67%</td>
<td>67%</td>
</tr>
</tbody>
</table>

Table 13. Accuracy in judgments about first and second themes in the modified perceptual task, where the harmony was removed.
Musical Scores Used in the Perceptual Experiment

The following examples are some of the musical themes (as deemed by Barlow and Morgenstern) used in the perceptual study. The themes are presented in their original keys, instead of transposed to C major or minor.

Figure 24. Beethoven’s *Piano Sonata No. 22 in F Op. 54, Mvt 1*. First theme (top) and second theme (bottom).
Figure 25. Brahms’ *Rhapsody in E-Flat Op. 119, No. 4*. First theme (top) and second theme (bottom).
Figure 26. Dvořák’s *Waltzes Op. 54, No. 3*. First theme (top) and second theme (bottom).
Figure 27. Prokofiev’s *Sonata No. 6 Op. 82, Mvt 1*. First theme (top) and second theme (bottom).
<table>
<thead>
<tr>
<th></th>
<th>Waltzes, Op 54, No 3</th>
<th>Sonata in D K 311, Mvt 1</th>
<th>Movements Perpetuels, No 3</th>
<th>Sonata No 6 Op 82, Mvt 1</th>
<th>Rhapsody in E Flat Op 119 No 4</th>
<th>Minuet in G</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Composer</strong></td>
<td>Dvorak</td>
<td>Mozart</td>
<td>Poulenc</td>
<td>Prokofiev</td>
<td>Brahms</td>
<td>Beethoven</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>1879</td>
<td>1777</td>
<td>1918</td>
<td>1939</td>
<td>1892</td>
<td>1805</td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>Expected</td>
<td>Expected</td>
<td>Expected</td>
<td>Expected</td>
<td>Expected</td>
<td>Expected</td>
</tr>
<tr>
<td><strong>Percent Correct</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>With DA</strong></td>
<td>20</td>
<td>35</td>
<td>50</td>
<td>33.3</td>
<td>87.5</td>
<td>54.2</td>
</tr>
<tr>
<td><strong>C.I.</strong></td>
<td>6</td>
<td>5.32</td>
<td>5.12</td>
<td>5.53</td>
<td>5.89</td>
<td>5.49</td>
</tr>
<tr>
<td><strong>Without DA</strong></td>
<td>16.7</td>
<td>66.7</td>
<td>79.2</td>
<td>25</td>
<td>90</td>
<td>65</td>
</tr>
<tr>
<td><strong>C.I.</strong></td>
<td>5.42</td>
<td>4.88</td>
<td>5.42</td>
<td>5.06</td>
<td>5.81</td>
<td>5.31</td>
</tr>
<tr>
<td><strong>Recordings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>With DA</strong></td>
<td>15</td>
<td>50</td>
<td>55</td>
<td>33.3</td>
<td>75</td>
<td>87.5</td>
</tr>
<tr>
<td><strong>C.I.</strong></td>
<td>5.37</td>
<td>5.04</td>
<td>4.99</td>
<td>6.4</td>
<td>5.94</td>
<td>6.24</td>
</tr>
<tr>
<td><strong>Without DA</strong></td>
<td>37.5</td>
<td>54.2</td>
<td>95.8</td>
<td>25</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td><strong>C.I.</strong></td>
<td>6.03</td>
<td>5.87</td>
<td>6.82</td>
<td>6.15</td>
<td>5.35</td>
<td>5.12</td>
</tr>
<tr>
<td><strong>Performance Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Dynamic</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>B Dynamic</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Dynamic Diff.</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>A Articulation</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>B Articulation</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Articulation Diff.</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Structural Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A mode</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B mode</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A key</td>
<td>E</td>
<td>D</td>
<td>F</td>
<td>A</td>
<td>E-</td>
<td>G</td>
</tr>
<tr>
<td>B key</td>
<td>D-</td>
<td>B-</td>
<td>F</td>
<td>C</td>
<td>A-</td>
<td>G</td>
</tr>
<tr>
<td>Key Distance</td>
<td>0.5</td>
<td>0.54</td>
<td>0</td>
<td>0.5</td>
<td>0.32</td>
<td>0</td>
</tr>
<tr>
<td>A pace</td>
<td>8</td>
<td>11.64</td>
<td>14.06</td>
<td>11.02</td>
<td>6.29</td>
<td>8.93</td>
</tr>
<tr>
<td>B pace</td>
<td>5.56</td>
<td>9.68</td>
<td>9.33</td>
<td>5.84</td>
<td>10.32</td>
<td>8</td>
</tr>
<tr>
<td>A nPVI</td>
<td>19.05</td>
<td>32.59</td>
<td>28.2</td>
<td>50.11</td>
<td>33.33</td>
<td>110.61</td>
</tr>
<tr>
<td>B nPVI</td>
<td>58.18</td>
<td>19.61</td>
<td>24.91</td>
<td>45.84</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>A interval size</td>
<td>1.71</td>
<td>2.05</td>
<td>1.87</td>
<td>1.83</td>
<td>1.95</td>
<td>1.86</td>
</tr>
<tr>
<td>B interval size</td>
<td>1.55</td>
<td>1.61</td>
<td>1.9</td>
<td>3.13</td>
<td>3.25</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 14. Participant scores and quantitative indices of musical works.
### Table 14 continued

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Composer</strong></td>
<td>Beethoven</td>
<td>Beethoven</td>
<td>Liszt</td>
<td>Mendelssohn</td>
<td>Falla</td>
<td>Chopin</td>
<td></td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>1804</td>
<td>1804</td>
<td>1881</td>
<td>1841</td>
<td>1902</td>
<td>1839</td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>Opposite</td>
<td>Opposite</td>
<td>Opposite</td>
<td>Opposite</td>
<td>Opposite</td>
<td>Opposite</td>
<td></td>
</tr>
<tr>
<td><strong>Percent Correct</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With DA</td>
<td>83.3</td>
<td>33.3</td>
<td>75</td>
<td>35</td>
<td>85</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>C.I.</td>
<td>5.9</td>
<td>5.87</td>
<td>5.97</td>
<td>5.61</td>
<td>4.96</td>
<td>5.82</td>
<td></td>
</tr>
<tr>
<td>Without DA</td>
<td>90</td>
<td>50</td>
<td>85</td>
<td>29.2</td>
<td>70.8</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>C.I.</td>
<td>6.1</td>
<td>4.74</td>
<td>5.39</td>
<td>5.47</td>
<td>4.97</td>
<td>5.51</td>
<td></td>
</tr>
<tr>
<td><strong>Recordings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With DA</td>
<td>79.2</td>
<td>41.2</td>
<td>66.7</td>
<td>30</td>
<td>85</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>C.I.</td>
<td>6.96</td>
<td>6.47</td>
<td>5.63</td>
<td>5.32</td>
<td>5.3</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>Without DA</td>
<td>100</td>
<td>40</td>
<td>45</td>
<td>25</td>
<td>54.2</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td>C.I.</td>
<td>5.98</td>
<td>5.62</td>
<td>4.79</td>
<td>5.88</td>
<td>5.3</td>
<td>5.14</td>
<td></td>
</tr>
<tr>
<td><strong>Performance Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Dynamic</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>B Dynamic</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Dynamic Diff.</td>
<td>-3</td>
<td>-5</td>
<td>-3</td>
<td>-3</td>
<td>-1</td>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>A Articulation</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>B Articulation</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Articulation Diff.</td>
<td></td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>-2</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td><strong>Structural Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A mode</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B mode</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>A key</td>
<td>F</td>
<td>f</td>
<td>F#</td>
<td>a</td>
<td>D</td>
<td>c#</td>
<td></td>
</tr>
<tr>
<td>B key</td>
<td>F</td>
<td>f</td>
<td>C#</td>
<td>a</td>
<td>A</td>
<td>c#</td>
<td></td>
</tr>
<tr>
<td>Key Distance</td>
<td>0</td>
<td>0</td>
<td>0.32</td>
<td>0</td>
<td>0.32</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>A pace</td>
<td>8.48</td>
<td>16</td>
<td>6.42</td>
<td>14.61</td>
<td>9.42</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>B pace</td>
<td>12</td>
<td>7.33</td>
<td>4.47</td>
<td>6.51</td>
<td>13.3</td>
<td>3.92</td>
<td></td>
</tr>
<tr>
<td>A nPVI</td>
<td>120.71</td>
<td>0</td>
<td>47.94</td>
<td>30.3</td>
<td>40.65</td>
<td>56.72</td>
<td></td>
</tr>
<tr>
<td>B nPVI</td>
<td>0</td>
<td>7.06</td>
<td>32</td>
<td>73.98</td>
<td>31.3</td>
<td>19.17</td>
<td></td>
</tr>
<tr>
<td>A interval size</td>
<td>3.31</td>
<td>2.21</td>
<td>3.52</td>
<td>4.32</td>
<td>1.7</td>
<td>4.18</td>
<td></td>
</tr>
<tr>
<td>B interval size</td>
<td>1.96</td>
<td>3.24</td>
<td>2.06</td>
<td>1.94</td>
<td>1.23</td>
<td>2.13</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 14 continued

<table>
<thead>
<tr>
<th>Composer</th>
<th>Sonata in F K 332, Mvt 3</th>
<th>La Vie Parisienne, Act 2 Entr'acte</th>
<th>Sonata No 2, Mvt 2</th>
<th>Hungarian Rhapsody No 4 in E Flat</th>
<th>Reverie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>1781</td>
<td>1866</td>
<td>1936</td>
<td>1853</td>
<td>1890</td>
</tr>
<tr>
<td>Type</td>
<td>Conflicting Articulation</td>
<td>Conflicting Articulation</td>
<td>Conflicting Articulation</td>
<td>Conflicting Articulation</td>
<td></td>
</tr>
</tbody>
</table>

**Percent Correct Scores**

<table>
<thead>
<tr>
<th></th>
<th>With DA</th>
<th>C.I.</th>
<th>Without DA</th>
<th>C.I.</th>
<th>Recordings</th>
<th>C.I.</th>
<th>Performance Data</th>
<th>Structual Data</th>
<th>Continuing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With DA</td>
<td>35</td>
<td>5.3</td>
<td>16.7</td>
<td>5.86</td>
<td>40</td>
<td>5.59</td>
<td>A Dynamic</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>5.37</td>
<td>50</td>
<td>5.44</td>
<td>50</td>
<td>5.5</td>
<td>B Dynamic</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>4.98</td>
<td>75</td>
<td>5.41</td>
<td>50</td>
<td>5.5</td>
<td>Dynamic Diff.</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>79.2</td>
<td>5.48</td>
<td>85</td>
<td>6.41</td>
<td>79.2</td>
<td>5.8</td>
<td>A Articulation</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70.8</td>
<td>5.91</td>
<td>65</td>
<td>4.47</td>
<td>54.2</td>
<td>5.5</td>
<td>B Articulation</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Articulation Diff.</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 14 continued

<table>
<thead>
<tr>
<th></th>
<th>Cracovienne Fantastique Op 14 No 6</th>
<th>Sonata No 14 in C Sharp Minor, Op 27 No 2, Mvt 2</th>
<th>Sonata No 21 in C Op 53, Mvt 1</th>
<th>Sonata in B Flat Op 47 No 2, Mvt 3</th>
<th>Polonaises Op 40 No 1 in A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composer</td>
<td>Paderewski</td>
<td>Beethoven</td>
<td>Beethoven</td>
<td>Clementi</td>
<td>Chopin</td>
</tr>
<tr>
<td>Date</td>
<td>1887</td>
<td>1801</td>
<td>1803</td>
<td>1820</td>
<td>1838</td>
</tr>
<tr>
<td>Type</td>
<td>Conflicting Dynamics</td>
<td>Conflicting Dynamics</td>
<td>Conflicting Dynamics</td>
<td>Conflicting Dynamics</td>
<td>Conflicting Dynamics</td>
</tr>
</tbody>
</table>

**Percent Correct**

<table>
<thead>
<tr>
<th></th>
<th>With DA</th>
<th>75</th>
<th>70.8</th>
<th>58.3</th>
<th>45</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C.I.</td>
<td>5</td>
<td>5.42</td>
<td>6.99</td>
<td>5.36</td>
<td>4.63</td>
</tr>
<tr>
<td>Without DA</td>
<td>50</td>
<td>60</td>
<td>45</td>
<td>37.5</td>
<td>79.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C.I.</td>
<td>5.27</td>
<td>5.06</td>
<td>5.78</td>
<td>5.62</td>
<td>5.49</td>
</tr>
</tbody>
</table>

**Recordings**

<table>
<thead>
<tr>
<th></th>
<th>With DA</th>
<th>70.8</th>
<th>66.7</th>
<th>66.7</th>
<th>65</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C.I.</td>
<td>5.8</td>
<td>6.9</td>
<td>6.51</td>
<td>5.61</td>
<td>5.31</td>
</tr>
<tr>
<td>Without DA</td>
<td>40</td>
<td>60</td>
<td>35</td>
<td>75</td>
<td>66.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C.I.</td>
<td>4.95</td>
<td>5.34</td>
<td>5.34</td>
<td>6.05</td>
<td>6.94</td>
</tr>
</tbody>
</table>

**Performance Data**

<table>
<thead>
<tr>
<th></th>
<th>A Dynamic</th>
<th>2</th>
<th>3</th>
<th>2</th>
<th>3</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B Dynamic</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Dynamic Diff.</td>
<td>-5</td>
<td>-2</td>
<td>-1</td>
<td>-3</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>A Articulation</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Articulation</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Articulation Diff.</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Structural Data**

<table>
<thead>
<tr>
<th></th>
<th>A mode</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B mode</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A key</td>
<td>B</td>
<td>A-</td>
<td>C</td>
<td>B-</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B key</td>
<td>B</td>
<td>D-</td>
<td>F-</td>
<td>B-</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Key Distance</td>
<td>0</td>
<td>0.32</td>
<td>0</td>
<td>0</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>B pace</td>
<td>11.44</td>
<td>3.6</td>
<td>2.8</td>
<td>6.92</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>A nPVI</td>
<td>19.61</td>
<td>62.22</td>
<td>31.89</td>
<td>13.33</td>
<td>51.11</td>
<td></td>
</tr>
<tr>
<td>B nPVI</td>
<td>38.01</td>
<td>24.1</td>
<td>29.63</td>
<td>29.82</td>
<td>77.95</td>
<td></td>
</tr>
<tr>
<td>A interval size</td>
<td>1.41</td>
<td>2.67</td>
<td>1.96</td>
<td>2.15</td>
<td>2.73</td>
<td></td>
</tr>
<tr>
<td>B interval size</td>
<td>2.47</td>
<td>3.68</td>
<td>1.78</td>
<td>1.7</td>
<td>5.67</td>
<td></td>
</tr>
</tbody>
</table>