9-2024

A Multilevel Approach to the Analysis and Visualization of Timbral Brightness in Post-Tonal Music

Stephen Spencer
The Graduate Center, City University of New York

A MULTILEVEL APPROACH TO THE ANALYSIS AND VISUALIZATION OF TIMBRAL BRIGHTNESS IN POST-TONAL MUSIC

by

STEPHEN SPENCER

A dissertation submitted to the Graduate Faculty in Music in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2024

Approval

A Multilevel Approach to the Analysis and Visualization of Timbral Brightness in Post-Tonal Music

by

Stephen Spencer

This manuscript has been read and accepted for the Graduate Faculty in Music in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Approved: August 2024

Yayoi Uno Everett, Chair of Examining Committee
Norman Carey, Executive Officer

Supervisory Committee:

Johanna Devaney, advisor
Joseph Straus, first reader
Yayoi Uno Everett
Robert Hasegawa

CITY UNIVERSITY OF NEW YORK

Abstract

A Multilevel Approach to the Analysis and Visualization of Timbral Brightness in Post-Tonal Music

by

Stephen Spencer

Advisor: Johanna Devaney

This dissertation develops a method for the analysis and visualization of timbral brightness in post-tonal music. Brightness, a salient timbral attribute correlated with the prominence of high frequencies in the sound spectrum, is crucial for distinguishing sounds and discerning how they fit together in music. The dissertation employs a multilevel analytical approach, measuring differences in brightness at multiple "textural levels" and temporal scales. The analytical utility of the approach is demonstrated through four case studies, each one closely examining the score and audio of a piece of instrumental music written in the first quarter of the twentieth century.

Chapter 1 defines three textural levels—the level of the *element, layer,* and *texture*—and outlines the methodology using a short excerpt from Igor Stravinsky's *Rite of Spring*. The next three chapters (2–4) each focus on a single textural level: Chapter 2 foregrounds *element*-level brightness relationships in Arnold Schoenberg's "Farben," Chapter 3 examines the relative brightness of polyphonic *layers* in Ruth Crawford's *Music for Small Orchestra*, and Chapter 4 analyzes the overall brightness of complete *textures* in three recordings of Anton Webern's *Six Pieces for Orchestra,* Op. 6/II.

The dissertation culminates in Chapter 5 with an analysis of the first movement of Edgard Varèse's *Octandre* that integrates all three textural levels. Together, these five chapters demonstrate a comprehensive yet flexible approach, bridging the gap between score and audio and pointing a way forward for the analysis of a salient but neglected feature of music.

Acknowledgments

To my advisor, Johanna Devaney: thank you for your unwavering support and guidance in the development of this dissertation. Your far-reaching advice, perceptive comments, and careful editorial approach enriched this project immeasurably and helped me grow as a scholar and teacher. To my first reader, Joe Straus, and committee members Yayoi Uno Everett and Bob Hasegawa: thank you for your incisive comments and encouraging feedback. Your input was instrumental in shaping this dissertation and will continue to impact my future work.

To my mentors at CUNY—Suzanne Farrin, Mark Spicer, Jason Eckardt, Poundie Burstein, Jeff Nichols, and Bill Rothstein: thank you for your inspiring seminars and conversations. I am so grateful to have learned from and with you. To my friends and colleagues at CUNY—Matt Sandahl, Jake Sachs-Mishalanie, Charles Rudig, Eliav Kol, Charles Colwell, Erin Johnston, Stephen Gomez, Lina Tabak, Nathan Pell, and countless others: thank you for your friendship and solidarity over the years, and for creating the warmest and most supportive academic environment I've ever been a part of. To Moe Touizrar and Dan Garmon: thank you for our invigorating conversations about music and life, the impact of which cannot be overstated. To Eric West and Branden Goodman: thank you for decades of friendship and laughter.

To my parents, Margaret and Mendel Spencer: thank you for your boundless love and encouragement throughout my life, and for the profound gift that was my music education. To my wife, Angela, for whom I do everything: thank you for your love and support throughout this process, for always asking the most important questions, and for always making me feel understood. To my daughter, Rose: thank you for radiating joy and wonder at the world, and for inviting me to see it through your eyes.

Contents

Figures viii
Tables xiii

Introduction 1
Overview of Chapters 3

Chapter 1. Analyzing Timbral Brightness: A Multilevel Approach 5
Capturing Brightness via the Spectral Centroid (SC) 5
Elements, Layers, and Textures 11
Conclusion 23

Chapter 2. Chord Color as Brightness in Arnold Schoenberg's "Farben" 25
Existing Approaches to "Farben" 25
Chord Colors and their Composition 31
Colors in Time 39
Sequential Element Contours and Permutations 49
Sketch of a Layer-Level Analysis 54
Conclusion 58

Chapter 3. Visualizing the Relative Brightness of Concurrent Textural Layers in Ruth Crawford's *Music for Small Orchestra* 59
Partitional Approaches to Texture 59
Textural Partitions in *Music for Small Orchestra* 63
Quantifying Coordination: Agglomeration and Dispersion 65
Interlayer Brightness 71
Clustering and Transparency 76
Intralayer Brightness 81
Conclusion 85

Chapter 4. Texture-Level Brightness and Volatility in Anton Webern's *Six Pieces for Orchestra*, Op. 6/II 86
Segments, Blocks, and Sections in Op. 6/II 87
Texture-Level Brightness Analysis 92
Conclusion 107

Chapter 5. Timbre as an Agent of Delineation in Edgard Varèse's *Octandre* 109
Varèse on Timbre 109
Levels of Timbral Design 114
Octandre, I. mm. 1–9 117
Octandre, I. mm. 10–16 122
Octandre, I. mm. 19–25 131
Octandre, I. mm. 26–32 137
Timbre as an Agent of Delineation 140

Conclusion 146
Future Directions 148

Appendix: Annotated Score of *Music for Small Orchestra* (excerpt) 155
References 157

Figures

1-1. (a) Raw spectral centroid data for clarinet; (b) same for cello; (c) average SC for each instrumental tone with SCQ labeled (2.17) 7

1-2. (a) Contrast in brightness despite continuity in pitch/loudness; (b) continuity of brightness despite contrast of pitch/loudness 11

1-3. (a) Elements in Stravinsky's *Rite of Spring*; (b) Elements plotted as points 13

1-4. Analyzing a verticality via concurrent element-level centroids in Stravinsky's *Rite of Spring*. (a) Score reduction. (b) Centroid plot 14

1-5. Analyzing a succession of events via sequential element-level centroids in Stravinsky's *Rite of Spring*. (a) Score reduction. (b) Centroid plot 15

1-6. (a) Layers in Stravinsky's *Rite of Spring*; (b) Layers plotted as points 17

1-7. (a) Beat-level slices of the complete texture in Stravinsky's *Rite of Spring* (Intro to Part I, R4); (b) Centroid of each slice in three recordings; (c) Centroid of each slice averaged across recordings 19

1-8. (a) Reduced score incipits of three adjacent sections in Stravinsky's *Rite of Spring* (R3–5); (b) Texture-level centroid of those sections in three recordings; (c) Texture-level centroid averaged across recordings 22

2-1. Engraved segment of Charles Burkhart's Fig. 9 (1973, p. 159), top four staves only. The series of combinations numbered 1–11 is repeated in retrograde. 26

2-2. Concurrent element graph of the opening two sonorities in "Farben." The first chord is evenly spaced, while the second is highly clustered 32

2-3. Changing ordering of the voices in the "Farben" chords (labeled 1 and 2), showing the alto voice and low tenor voice shifting to the top of the sonority. 37

2-4. Annotated concurrent layer graph showing unevenness, span, voice ordering, and clustering. 38

2-5. Concurrent element graph showing color differences between the repeating opening sonorities and the string punctuation in measure 11 41

2-6. The first three instrument combinations in "chord color space," showing average element-level centroid plotted against span 42

2-7. (a) Concurrent element graph of four sonorities in mm. 13–14. (b) The same four sonorities in chord color space 44

2-8. Sonorities 8–24 (mm. 15–23) in chord color space, with two opposing three-chord segments labeled 46

2-9. “Inverted” span changes, m. 32 48

2-10. Sequential element graph of mm. 13–19 of “Farben” 50

2-11. A seven-tone sequence of element-level centroids returns in the final section of the work in retrograde. 51

2-12. (a) The entire tenor 2 brightness contour from mm. 13–18 recurs in retrograde inversion in the tenor 1 voice in mm. 33–38. (b) Segments of the tenor 1 part from mm. 13–19 appear in retrograde inversion in the tenor 2 part in mm. 32–38 53

2-13. Concurrent layer graph of mm. 15-25. The organism’s overall centroid inches downward in tandem with pitch 55

2-14. In mm. 15–25, the foreground material becomes brighter as the organism becomes darker 56

3-1. Three configurations of six elements in Wallace Berry’s textural notation. 61

3-2. Partitional score annotation of mm. 3–9 of Crawford’s *Music for Small Orchestra* 63

3-3. Partitional reading of mm. 3–9, revised to account for unison doublings 65

3-4. Coordination graph of mm. 3–9 of *Music for Small Orchestra*, showing an alternation between agglomerative (bottom) and dispersive (top) tendencies 67

3-5. Accumulating dispersion in mm. 1–44 of *Music for Small Orchestra*. 69

3-6. (a) Interlayer similarity in mm. 10–14 of *Music for Small Orchestra*; (b) Intralayer dissimilarity in mm. 56–59 70

3-7. (a) Centroid data; (b) SCQ data; (c) Resulting brightness analysis of concurrent layers in the first textural unit (mm. 3–5) 73

3-8. Concurrent layer graphs of the first three textural units 75

3-9. (a) Textural Unit 6 groups into three clusters, while Textural Unit 7 groups into two. (b) Textural Unit 7 exhibits greater opacity and less transparency than Textural Unit 6. (c) Score fragments of each Textural Unit reordered according to brightness. 78

3-10. (a) Concurrent layer graph of mm. 3-44 of *Music for Small Orchestra*; (b) Normalized Transparency graph of the same passage; (c) Increasing opacity correlated with more layers and fewer clusters 80

3-11. Left: Element-level spectral centroids of the layers comprising Textural Unit 1, with unisons connected by square brackets. Right: Box and whisker plots summarizing the spread of the six-part chord in its textural context 82

3-12. Intra- and interlayer brightness relationships of the gong chord and cello ostinato. (a) The gong chord contains a bright cello cluster and a dark piano cluster, connected by unisons. (b) The first two notes of the cello ostinato fit into the gap between the gong-chord clusters. (c) Box and whisker plot summarizing layers..... 84

4-1. Instrumentation graph and formal segmentation of Webern's Op. 6/II. Boundaries at three temporal levels (segment, block, and section) align with changes in orchestration. 88

4-2. My formal segmentation aligned with Baker's (1982) reading of the form of Op. 6/II. Most of Baker's units align with blocks, but some align only with "segments." 91

4-3. (a) Mean centroid ("sectional average") of formal segments of Op. 6/II in three performances. (b) Mean centroid with standard deviation. 94

4-4. (a) Sectional average for segments i–vi across three recordings, showing upward trend in brightness. (b) SCQs for each adjacent segment-pair, outliers in bold....... 96

4-5. Mean centroid (a) and standard deviation (b) for three recordings of Webern Op. 6/II, formal segments v–vii 97

4-6. Reduced score of segments vi and vii. In segment vi, the crescendo in all instruments in leads to wide deviations in spectral centroid (cf. Figure 4-5b)......... 98

4-7. Mean centroid of each segment in all three recordings, broken down by section (Section D excluded). Mean and standard deviation given for each section per recording 99

4-8. Sectional average for segments vii–xiii across three recordings, showing mean and standard deviation of the section 100

4-9. Average spectral centroid and pitch centroid of Section C. Dotted rectangles show peaks in pitch (corresponding to the "steam whistle" motive), solid rectangle shows peaks in brightness (muted brass tutti) 102

4-10. Subsegments from mm. 24–27 across three recordings. Average SCQs between subsegments reported 104

4-11. (a) Segment centroids averaged across all three recordings. Bands show the spread of recordings. Dotted line shows increasing brightness trend. (b) Segment centroids broken down by section 105

5-1. Overview of levels analyzed in each section of first movement of *Octandre*, aligned with instrumentation graph. 116

5-2. (a) Segmentation of oboe solo roughly by note; (b) average brightness of each note across five recordings, with standard deviation of recordings. (c) oboe solo notes broken down by recording, with average note-level centroid and standard deviation given. 119

5-3. Average centroid of the solos that open each movement of *Octandre* 120

5-4. (a) The texture-level centroid darkens at the entry of the clarinet, and stays at roughly the same level when the bass joins. (b) Layer-level analysis of mm. 5–6, with SCQs labeled 122

5-5. (a) Reduced score of mm. 10–16 of *Octandre*, I., with formal segmentation and partitional analysis 124

5-6. (a) Layer-level centroids of mm. 10–16. Colors match points to layers in Figure 5-5. (b) Texture-level centroids of same passage 125

5-7. Increasing agglomeration trend for *Octandre*, I., up to m. 20: textural layers are getting larger 128

5-8. (a) Layer-level analysis showing outward expansion from a duo layer to two trio layers. (b) Element-level analysis showing narrow wind trio nested within dispersed brass trio. 129

5-9. (a) Score reduction and segmentation of mm. 19–22. (b) Texture-level analysis (overall centroid) of segments in mm. 19–22 of *Octandre*; (c) Texture-level analysis (overall centroid) of subsegments of same passage 132

5-10. Increased playing intensity in m. 21 correlates with increased texture-level brightness (Avg. SCQ = 2.91) 133

5-11. Element-level analysis of mm. 19–22 135

5-12. Element-level analysis of mm. 26–29 138

5-13. Texture-level analysis, mm. 26 to the end of the piece 140

5-14. Multilevel analysis of Octandre, I. (a) Texture-level centroids of entire movement; (b) Instrumentation of the piece; (c) Layer-level analysis of mm. 5–18; (d) Element-level analysis of mm. 17–29 141

5-15. Average texture-level brightness of the first two sections. 142

Tables

2-1. (a) Unevenness scores for Chords 1 and 2 in “Farben.” (b) Silhouette scores of *k*-means clustering. Element-level centroids in Chord 1 do not appear to be strongly clustered, while elements in Chord 2 appear to form three clusters 31

Introduction

Timbre encompasses a vast set of perceptual attributes that we use to identify, discriminate, and compare sounds. A prevalent approach to analyzing timbre in music, amidst the diverse array of strategies that have proliferated in recent years, involves quantifying these perceptual attributes through acoustic analysis: measuring features of audio signals that correlate with our auditory perceptions. Given the multitude of attributes that fall under the umbrella of timbre, music theorists interested in timbral analysis must navigate an overwhelming number of acoustic features, which overlap with each other and give rise to our listening experiences in complicated ways.

The *brightness* of a sound, correlated with the presence and prominence of high frequencies in the spectrum, stands out as an especially salient timbral attribute. Listeners rely heavily on brightness to distinguish sounds and discern how they fit together—whether they merge into coherent units or split into distinct elements—as they overlap and interact in time. Its relatively straightforward acoustic measurement affords a wide range of analytical utilities that have not yet been explored systematically by music theorists, making it an important area for thorough investigation.

This dissertation focuses on timbral brightness, developing a systematic method for its analysis and visualization. While brightness is pertinent to many kinds of music, I show that it is particularly valuable for discussing the aesthetic and structural nuances of post-tonal instrumental music, a repertoire whose sonic complexity remains underserved by pitch-based methods. In the five chapters that follow, I develop a "multilevel" analytical approach, attempting to capture the multiple ways that brightness gives rise to our perceptions of post-tonal polyphony. I explore the analytical affordances and limitations of the approach through four case studies, each one closely examining the score and audio of a piece of instrumental music written in the first quarter of the twentieth century.

In polyphonic music, brightness operates at many scales of detail: we can speak of the brightness of individual tones, of certain combinations of tones, or of the complete polyphonic totality. Methodologies that confine themselves to a single scale of detail thus neglect critical aspects of timbral design. The central approach of the dissertation is to measure brightness at multiple "textural levels"—from extracted musical *elements*, to overlapping polyphonic *layers* and the overall *texture*. Analysis at these multiple levels not only affords comprehensiveness, but also flexibility: analysts may choose to focus at one textural level or observe interactions between multiple levels depending on the specific musical context.

Brightness perception is inherently relative and influenced by prior context (Siedenburg et al, 2021). Assessing the absolute brightness of an isolated sound therefore overlooks the broader perceptual effect of the sound in its musical surroundings. My approach in this dissertation is to emphasize the *relative* brightness of two or more adjacent sounds. I develop a ratio-based system of measurement that allows me to orient my analyses toward differences, distances, and transformations of sounds, rather than their absolute, context-independent qualities. Throughout the dissertation, I make these brightness comparisons at multiple "temporal scales," from short snippets of music to larger segments and entire formal sections.

The dissertation thus offers a method for evaluating via acoustics the relative timbral brightness of sounds at multiple textural levels and temporal scales. As mentioned, a primary goal of the dissertation is to capture through acoustic analysis the role that timbral brightness plays in shaping our perceptions of post-tonal polyphonic structures. But the dissertation also shows that attending more closely to brightness, as quantified by acoustic measurement, can bring out surprising spectral nuances that previously eluded our notice. The analyses in Chapters 2–5 invite us to reconcile what we hear—and what we see in the score—with quantitative measurements of brightness, and in so doing orient our listening toward the relative spectral position of sounds. Through this process, we are prompted to revise our initial assumptions about the music and challenge our well-worn habits of listening.

Overview of Chapters

In Chapter 1, I outline my multilevel methodology. In dialogue with a short excerpt from Igor Stravinsky's *Rite of Spring* (1913), I explain how to measure and visualize brightness at each of three textural levels—*element*, *layer*, and *texture*—and explore the kinds of insights each level affords. Each of the next three chapters (Chapters 2–4) focuses on a single textural level. Chapter 2 foregrounds *element*-level brightness in a discussion of the changing orchestral colors of Arnold Schoenberg's "Farben" (1909). Responding to preexisting timbral analyses of the piece that focus on either instrumentation or spectrogram analysis, I explore brightness relationships between the five elements of the opening chord and trace their transformation through the piece, revealing logical patterns that otherwise go undetected. Chapter 3 examines the relative brightness of concurrent textural *layers* in Ruth Crawford's *Music for Small Orchestra* (1926). The chapter develops a method for analyzing textural "transparency," problematizing existing approaches to texture that privilege rhythmic coordination. In Chapter 4, I zoom out to trace the overall brightness of complete *textures* in three recorded performances of Anton Webern's *Six Pieces for Orchestra,* Op. 6/II (1909), considering how large-scale brightness trajectories work to communicate the poetic concerns of the piece.

The dissertation culminates in Chapter 5 with an analysis of the first movement of Edgard Varèse's *Octandre* (1923) that integrates multiple textural levels. The chapter revolves around Varèse's conception of timbre "as an agent of delineation," employing multilevel brightness analysis to examine timbre's role in the form of the work. Together, these five chapters demonstrate an adaptable multilevel approach that bridges the gap between score and audio and points a way forward for the analysis of a perceptually salient, but analytically neglected feature of post-tonal music. The dissertation concludes with a discussion of the methodology's applicability to other musical repertoires and a consideration of future directions.

Chapter 1

Analyzing Timbral Brightness: A Multilevel Approach

This chapter details my methodology for analyzing and visualizing brightness in post-tonal music. In dialogue with empirical literature, I begin by describing the correlation between a sound's brightness and its spectral centroid—the midpoint of its spectral energy distribution. I propose a simple ratio, the spectral centroid quotient (SCQ), to quantify relations of brightness between two or more sounds, which I argue aligns with empirical research on brightness perception. Then, I define three "textural levels" for measuring brightness in post-tonal music: the level of the *element*, *layer*, and *texture*. With reference to a brief excerpt from Stravinsky's *Rite of Spring* (1913), I describe my technique for analyzing and visualizing brightness at each level, highlighting the insights each textural level affords.

Capturing Brightness via the Spectral Centroid (SC)

The perception of brightness, like all auditory attributes, depends on acoustic properties. Scholars have frequently noted, since at least the work of Helmholtz (1863/1875), that brightness correlates with the presence and prominence of high frequencies in the sound spectrum. This can be quantified using a simple statistical approach first proposed by Lichte (1941), who defines brightness as "a function of the location [in Hz] of the midpoint of the energy distribution" (472). The now-standard term for this midpoint—the *spectral* *centroid* (SC)—was first used by Grey and Gordon (1978), who compare it to the "center of gravity" of the frequency spectrum.[1]

In modern audio signal processing, the SC is calculated by taking a weighted average of all the frequencies present in a spectrographic representation of an audio signal, with the magnitudes as the weights:

$$\mathrm{SC} = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

where *x(n)* represents the magnitude of the frequency bin number *n,* and *f(n)* represents the center frequency of that bin. The SC is calculated as above for every audio frame in the signal, resulting in a long series of centroids.[2] Figure 1-1 plots the frame-level SC data of two instrumental tones, clarinet (1-1a) and cello (1-1b), each producing the pitch A4 at a mezzo-forte dynamic for one second. The prominence of higher frequencies in the cello spectrum (largely brought about by the noisiness of the bow sound) yields a higher overall SC.

---

[1] The strong correlation between the SC and perceived brightness is confirmed by numerous studies that collect judgments of brightness and compare the ordering of the data to the SC. While these experiments consistently yield one-dimensional brightness scales that highly correlate to the SC, there is nevertheless a slight tension between "perceived brightness," which varies between listeners, and what we might call "acoustic brightness" which is perfectly modelled by the SC (Schubert and Wolfe, 2006; Zacharakis et al, 2014; Almeida et al, 2017; Saitis and Siedenburg, 2020).

[2] In audio signal processing, a frequency *bin* is a discrete range of the frequency domain (i.e., one of many narrow, equal-sized slices of a spectrum) used for analysis and manipulation of signals. An *audio frame* is a small temporal segment of audio, typically a few milliseconds long. Its size is determined by the audio sample rate.

For the analytical purposes of this dissertation, the raw SC data shown in Figure 1-1a and 1-1b provides an excessive amount of detail; I am interested in visualizing the overall brightness level of discrete musical events, and less so in their fine-scale fluctuations at the level of the audio frame. Thus, I calculate a mean centroid for each sound by dividing the sum of all frame-level centroids by the total number of audio frames. Figure 1-1c shows the results of that averaging procedure for the same two instrumental sounds, plotting the results as two points on a single graph. A complex sound spectrum, changing through time, is thus summarized by a single mean value that represents its overall brightness.[3] While subtle particulars are reduced out in both the frequency and temporal domain, this reduction greatly simplifies the broad spectral comparison of sounds.

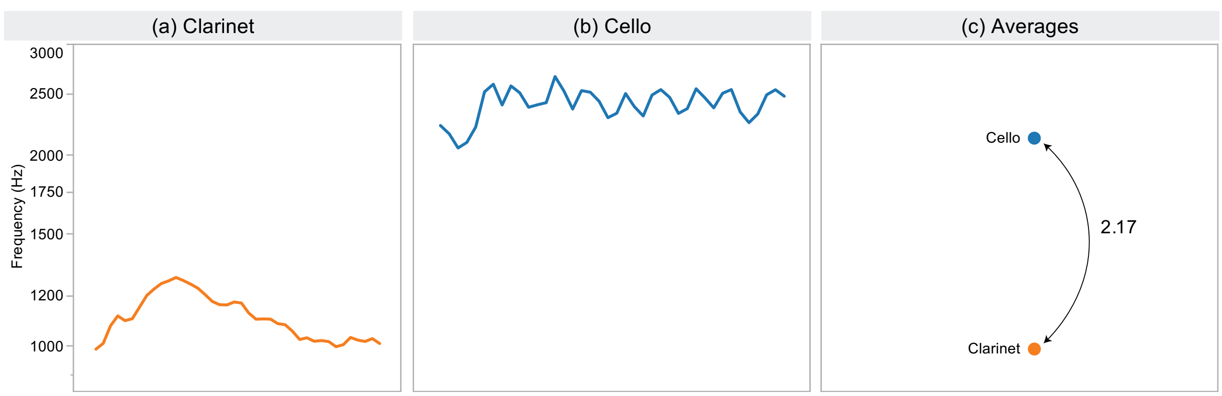


Figure 1-1. (a) Raw spectral centroid data for clarinet; (b) same for cello; (c) average SC for each instrumental tone, with SCQ labeled (2.17). The cello is roughly twice as bright as the clarinet. Audio samples taken from Garritan Personal Orchestra 5.

[3] After calculating the mean, it may be fruitful to reconsider the range of the fine-scale fluctuations, which can be represented by the standard deviation of centroids (i.e., the spread of frame-level centroids around the mean centroid). I do this in Chapter 4 when longer spans of audio are analyzed.

Measuring the *relative brightness* of musical events is a central preoccupation of the dissertation. In the chapters that follow, I quantify relative brightness of pairs of sounds using a simple ratio I call the spectral centroid quotient (SCQ). This is the value that results when we take the higher SC (*a*) of the pair and divide it by the lower one (*b*):

$$\mathrm{SCQ} = \frac{\mathrm{SC}_a}{\mathrm{SC}_b}$$

This ratio serves as a measure of the relative brightness of sounds *a* and *b*, specifically indicating how many times brighter *a* is in comparison to *b*. Figure 1-1c labels the SCQ (2.17) between the clarinet and cello timbres—the cello's A4 is thus roughly *twice as bright* as the clarinet's. I use ratios here, as opposed to absolute differences of centroid, to account for the logarithmic perception of frequency: a given SCQ will remain perceptually consistent across different frequency ranges, whereas absolute differences indicate varying degrees of perceived contrast depending on the starting frequency.

My emphasis on contrast over absolute brightness is supported by recent empirical research. Firstly, research has suggested that listeners can indeed perceive brightness relationships as ratios, just as the SCQ represents them (Almeida, 2017; Kazazis et al, 2021, 2022).[4] Perhaps more crucially, though, empirical findings suggest that auditory brightness perception is inherently relative and sensitive to prior context. Just as our eyes

[4] Almeida et al (2017) asked participants to adjust a synthesized sound to make it "twice as bright" as a reference sound. They found that the ratio of SCs required to double the brightness was around 2.00 on average, suggesting that brightness may be a dimension that, like pitch, forms a ratio scale. This finding is confirmed in a recent set of studies (Kazazis et al, 2021; 2022).

need time to adjust when we enter a dark room and turn on the light, so too do sounds appear brighter when preceded by darker sounds (Siedenburg, 2018; Siedenburg at al, 2021). My relative approach thus encourages an orientation toward perceptual contrast between adjacent sounds, allowing us to remain sensitive to contextual effects, rather than using the spectral centroid to make absolute claims about the brightness of each individual sound. This approach thus stands in contrast to "oppositional" approaches to timbre analysis (e.g., Cogan, 1984; Lavengood, 2020), wherein analysts categorize sounds as either holistically bright or dark based on spectral centroid measurements. Such binary analyses are insufficient insofar as they overlook the contextual and adaptive nature of brightness perception.[5]

Throughout the dissertation, centroid ratios (SCQs) are used to support claims about (a) degrees of either continuity or contrast between sequential sounds, and similarly, (b) degrees of either blend or segregation between concurrent sounds. Each of these will be examined in turn. Perceptual studies demonstrate that listeners rely heavily on brightness to make judgments about timbral similarity when two sounds of equal pitch and loudness are presented in succession (Grey and Gordon, 1978; Krimphoff et al, 1994; McAdams et al, 1995; Lakatos, 2000; Caclin et al, 2005). Thus, for two sequentially presented tones of the same pitch, a small SCQ suggests perceived continuity, while a larger SCQ suggests

[5] This project is closer to Kristi Hardman's approach to timbre analysis, which focuses on *continua* rather than *oppositions* (Hardman, 2022), but differs in its reliance on score data and its emphasis on relations between pairs of discrete sounds.

perceived contrast. This generalization requires qualification when pitch is not normalized: as pitches and dynamics of the two tones diverge, the contrast resulting from pitch and loudness differences becomes more prominent and the influence of the spectral centroid weakens. Thus, a refined generalization would be that perceptual continuities resulting from pitch and loudness similarity are *enhanced* by brightness similarity (modeled by small SCQs) and *reduced* by brightness difference (large SCQs). Conversely, contrasts in pitch and loudness are weakened by brightness similarity and strengthened by brightness difference. Figure 1-2 clarifies this relationship, showing in the left panel (1-2a) a sequence of tones in the same pitch and dynamic range, but with very different SCs, and in the right panel (1-2b) a sequence of highly contrasting tones with identical SCs. [6]

Similarly, when two sounds are presented concurrently, research has indicated that brightness (via SC) predicts whether they will become perceptually integrated or else segregate into distinct auditory streams.[7] Thus, lower concurrent SCQs generally lead to more perceptual integration. This operates in concert with parameters typically implicated

---

[6] In this dissertation, I do not adjust my brightness measurements to account for differences in pitch. At least one study argues that brightness is better correlated with a "pitch-adjusted centroid," derived by dividing the SC by the fundamental frequency (Kendall and Carterette, 1996; cf. Schubert and Wolfe, 2006). I prefer to work from absolute centroids, embracing the entanglement of multiple parameters.

[7] Building on Bregman (1989), Hartmann and Johnson (1991), Gregory (1994) and Singh & Bregman (1997) each found that alternating tones could segregate into separate streams if subjected to contrasting spectral modifications. Sandell (1991, 1995) and Tardieu & McAdams (2012) found that perceptual blend between dyads was predicted by low SC differences and a low composite SC, indicating that sounds with similar brightness, especially if both are relatively dark, are more likely to blend. See also Fischer et al (2021).

in stream segregation—e.g., onset synchrony, harmonicity, and shared contour (Bregman, 1989)—serving to enhance or diminish the perceptual relatedness brought about by those parameters. For example, two sounds that are not rhythmically aligned nor related by a harmonic interval might nevertheless be perceptually integrated by virtue of brightness similarity.

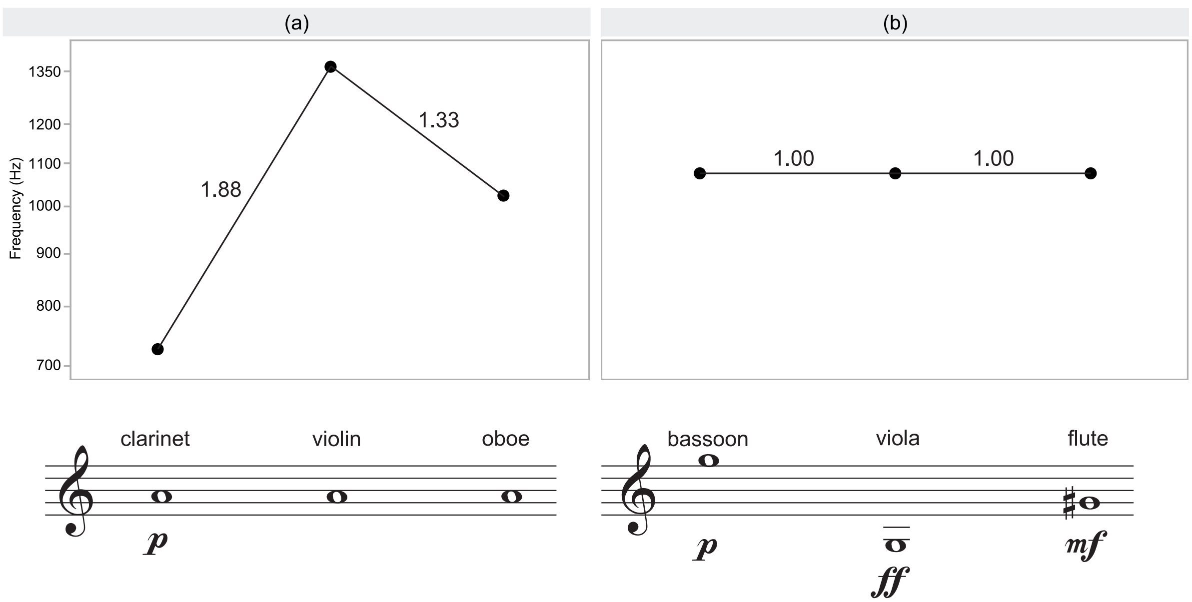


Figure 1-2. (a) contrast in brightness despite continuity in pitch/loudness; (b) continuity of brightness despite contrast of pitch/loudness. Audio samples taken from the Philharmonia Sample Library.

Elements, Layers, and Textures

Thus far my examples have featured isolated instrumental tones, but brightness operates at many levels; any sound spectrum, of any complexity and length, can be analyzed for its average spectral centroid. For the purposes of my analyses of post-tonal music, I define three basic "textural levels" at which centroids can be measured: I term these the level of the *element* (isolated musical tones), the level of the *layer* (separate strands of the musical texture) and the level of the *texture* (the entire audio signal). SCQs at each textural level can be measured either concurrently or sequentially, and at multiple temporal scales, ranging from short snippets to longer spans of time. In this section I define each textural level and my approach to its measurement and visualization, with illustrative examples from a brief excerpt of Stravinsky's *Rite of Spring* (Intro to Part I, R4).

Elements are the kinds of sounds we've analyzed thus far (in Figures 1-1 and 1-2): isolated musical tones or sequences of tones played by a single instrument or voice. At the level of the element, centroids are computed and compared for single notes or melodic lines without necessarily considering how those notes or lines are grouped perceptually with other simultaneous sounds. Figure 1-3a shows a snippet from the Introduction to Part I of The *Rite of Spring* (R4), with elements highlighted at multiple temporal scales—some at the scale of the single event, some encompassing several notes.

In Figure 1-3b, I plot the centroids of those highlighted elements. To isolate and extract element-level centroids, I rely on MIDI data realized by orchestral sample libraries, which allows me to isolate individual sound sources and evaluate them separately.[8] I follow a five-step process to examine element-level centroids: (1) define elements through score-based analysis; (2) input the musical data of the elements (pitch, duration, dynamic) into a music notation program; (3) export the audio playback of each element as performed by a sample

---

[8] At the time of this writing, extracting audio features of separate sound sources within a complex polyphonic audio signal is reliable for pitch analysis, but for timbre remains unfeasible for the level of spectral precision this dissertation requires. While it would be preferable to extract element-level data directly from musical performances, MIDI data realized by a sample library serves as a practical alternative.

library within the notation program; (4) create a trimmed audio file for each exported element (i.e., remove leading and trailing silences); (5) compute the spectral centroid of each audio file.[9] The resulting centroid values are plotted along a logarithmic frequency scale.[10]

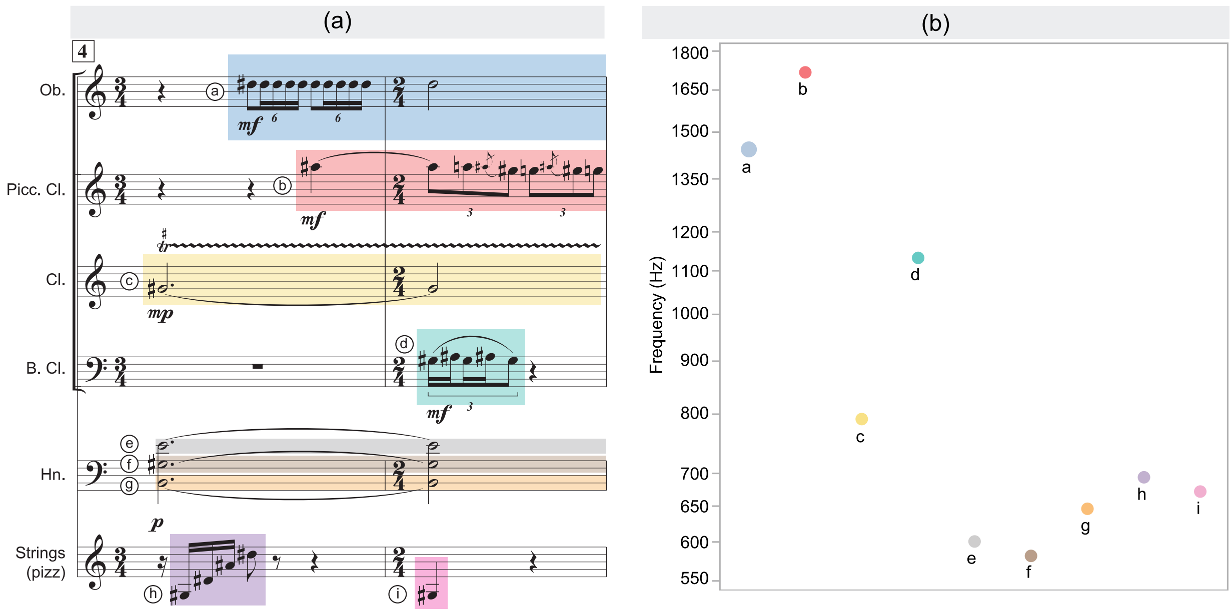


Figure 1-3. (a) Some elements in Stravinsky's *Rite of Spring* (Intro to Part I, R4); (b) those same elements plotted as points. Audio samples taken from Garritan Personal Orchestra 5.

Examining *concurrent* element-level centroids allows me to estimate the degree of integration of simultaneous instrumental tones, and thereby consider, as I do in Chapter 2, the perceptual "character" or "color" of verticalities in terms of their brightness make-up.

---

[9] All centroids in the dissertation are computed using the spectral centroid algorithm in the python library *librosa* (McFee et al, 2015).

[10] Throughout the dissertation, I adjust the range of the frequency axis for each plot to ensure that the centroids under consideration occupy the entire view. The visual distance between two centroids on the graph thus does not consistently reflect the magnitude of the SCQ between them.

This approach frames a chord as a vertical stack of element-level brightness values with a certain spread and distribution that contributes to the chord's overall character. This provides a new way to compare chords, offering a complement to traditional analysis of post-tonal harmony. The verticality that falls on the downbeat of the first measure in Figure 1-3—three horns and trilling clarinet—is plotted in Figure 1-4. The elements span a relatively low total region of the spectrum (outer SCQ=1.36), and on average sit rather low in brightness (avg. SC=657.6 Hz) relative to the rest of the texture (avg. SC=922.9 Hz). The horn tones, being as they are very similar in timbre, cluster together in brightness space, while the trilling clarinet is somewhat more isolated.

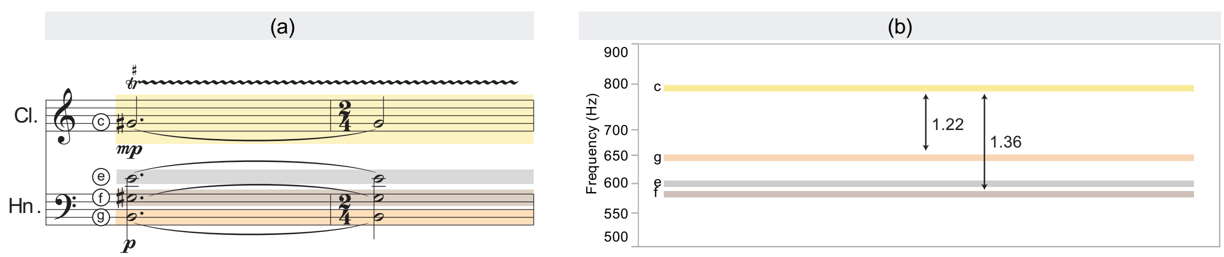


Figure 1-4. Analyzing a verticality via concurrent element-level centroids in Stravinsky's *Rite of Spring* (Intro to Part I, R4). (a) Score reduction. (b) Centroid plot. Audio samples from Garritan Personal Orchestra 5.

Element-level centroids can also be evaluated *sequentially*. This is useful for quantifying the relative brightness of successive elements—e.g., a melodic line that undergoes timbral modifications through time. Figure 1-5 picks out a brightness contour from the *Rite of Spring* example: the pizzicato violin timbre contrasts with the oboe (SCQ=2.08), which proceeds to the more closely related bass clarinet figure (SCQ=1.27). In Chapter 2, sequential element-level centroids allow me to compare several brightness contours in

Schoenberg's "Farben," tracing timbral changes through each voice of the piece's central polyphonic canon.

In my concurrent element-level visualizations of sustained chords, element centroids are shown as long horizontal lines indicating the duration of the corresponding tones (Figure 1-4). When considering collections of elements "out of time," (i.e., where temporal position and/or duration is less pertinent), I use circular points, as in Figure 1-3. Sequential element-level centroids—brightness contours—are always shown as points connected by lines, to emphasize visually their continuity through time (as in Figure 1-5).

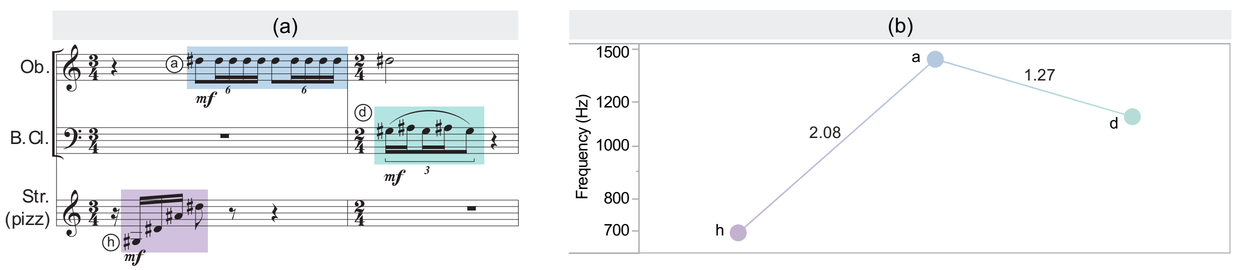


Figure 1-5. Analyzing a succession of events via sequential element-level centroids in Stravinsky's *Rite of Spring* (Intro to Part I, R4). (a) Score reduction. (b) Centroid plot. Audio samples from Garritan Personal Orchestra 5.

Layers are the separate strands of a musical texture. Analysis at the level of the layer requires us to consider which elements are grouped together into distinct units, either concurrently or sequentially, and then to compute the centroid of each unit as a whole. There are many conceivable criteria by which to segment a complex collection of musical elements into distinct layers, and different kinds of music will warrant different strategies. For the post-tonal music under consideration here, I group concurrent elements into layers primarily based on rhythmic coordination, and sequential elements into layers according to registral and gestural continuity. Thus, if the onsets of two or more elements are synchronized and remain coordinated for some time, I consider these elements together as a unit and plot a single centroid that represents their wholesale brightness. Similarly, if two sequential elements occupy the same register and/or exhibit similar musical behavior, I am likely to concatenate them into a single layer.[11]

The horns and clarinet, which enter together in the first measure of rehearsal no. 4, can be considered a single layer (shown in yellow and labeled “a” in Figure 1-6). Similarly, the bass clarinet gesture in the second measure is doubled by pizzicato violin at the unison for its first note, justifying their representation with a single layer-level centroid (shown in green and labeled “d” in Figure 1-6). There is a gestural continuity between the pizzicato violin in the first measure and the oboe’s repeated D♯5: the former ascends into and overlaps with the latter, prompting me to group them sequentially into a single layer (shown in blue and labeled “c” in Figure 1-6). The clarinet melody layer, containing just one element, is shown in red and labeled “b” in Figure 1-6.

[11] As mentioned, alternative segmentation methods may be more appropriate for different repertoires. For recorded popular music, for example, it would seem fruitful to group elements into layers based on textural function (Moore, 2012)—e.g., melody, harmonic filler, explicit beat layer, bass—and then compute the centroid for each. For music clearly stratified into melody and accompaniment, a figure vs. ground categorization (or foreground, middle ground, background) may be appropriate to determine which elements belong together.

Figure 1-6b illustrates the spectral layout of these simultaneous musical layers.[12] This layout captures the degree to which layers are perceptually integrated with each other. Generally, two layers separated by large ratios will remain perceptually distinct. For example, layers "a" and "b" in Figure 1-6b are separated by an SCQ of 2.66—the perceptual segregation brought about by rhythmic independence is thus *enhanced* by the stark difference in brightness. The two layers in the middle ("c" and "d") serve as a kind of spectral middle ground between the outer layers and are themselves closely related in brightness (SCQ=1.22)—here, the rhythm-based segregation is *weakened* by brightness similarity.

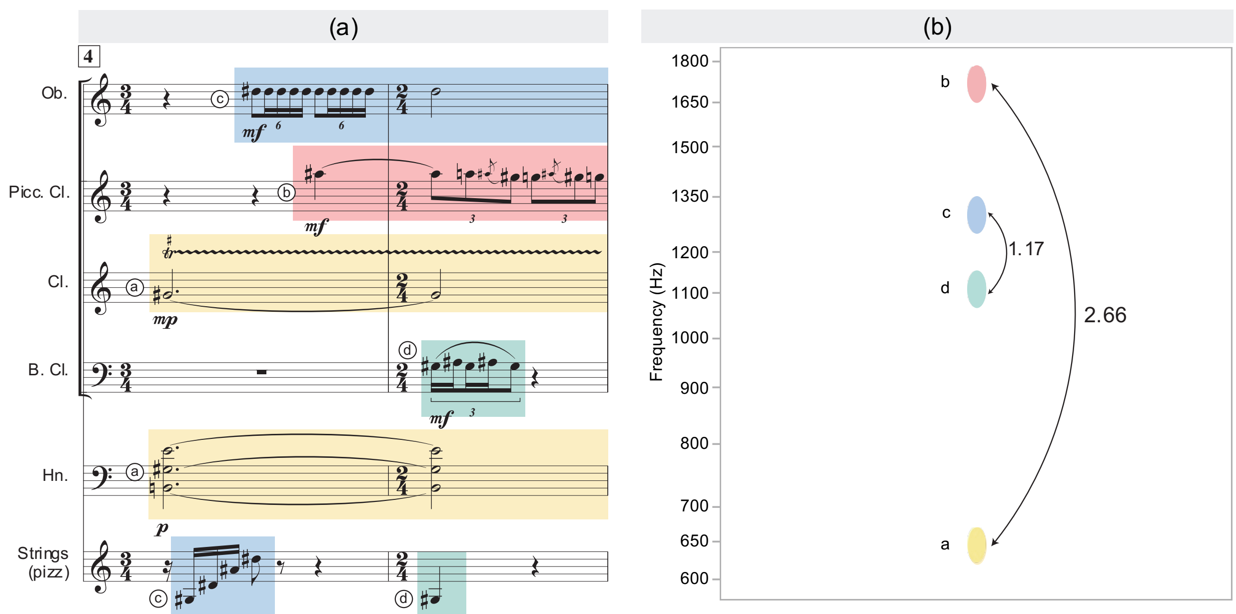


Figure 1-6. (a) Some layers in Stravinsky's *Rite of Spring* (Intro to Part I, R4); (b) Those same layers plotted as points. Audio samples taken from Garritan Personal Orchestra 5.

[12] To calculate these centroids, I follow the very same steps as at the level of the element, with one difference: when multiple elements cohere into a layer, I use the sample library to play them back simultaneously and export them as a single audio file.

To visually distinguish layer-level centroids from elements, my strategy is to use elliptical points. These points appear larger than the circular points used in element-level visualizations, reflecting the fact that layers often encompass multiple elements. And they direct the eye toward the vertical, subtly emphasizing vertical relationships between concurrent layers in the analysis. In these layer visualizations, colors are typically used to match centroids to an annotated score, but where possible this information is rendered redundant by the presence of other more accessible labels.

What I call the level of the *texture* is the entire audio signal, often encompassing multiple layers and elements. A texture-level centroid is a single value that represents the overall brightness of an all-encompassing vertical slice or segment of the score.[13] In Figure 1-7a, I highlight several such vertical slices—one per beat—in the *Rite of Spring* example.

Unlike elements and layers, my analyses at the level of the texture engage with audio recordings of actual performances. My approach is to analyze multiple performances of the same music, which allows me to assess an approximate range of centroids the music gives rise to. To demonstrate, Figure 1-7b plots a texture-level centroid for each the five beats highlighted in 1-7a; three recordings are analyzed and plotted separately using

[13] The global sound of the overall texture is equivalent to what some researchers have termed "polyphonic timbre" (e.g., Alluri & Toiviainen, 2010).

different shapes and colors.[14] I follow a four-step process here: (1) define textures through score-based analysis; (2) split audio recordings into segments corresponding to the defined textures; (3) export the segments as audio files; (4) compute the spectral centroid of each audio file.

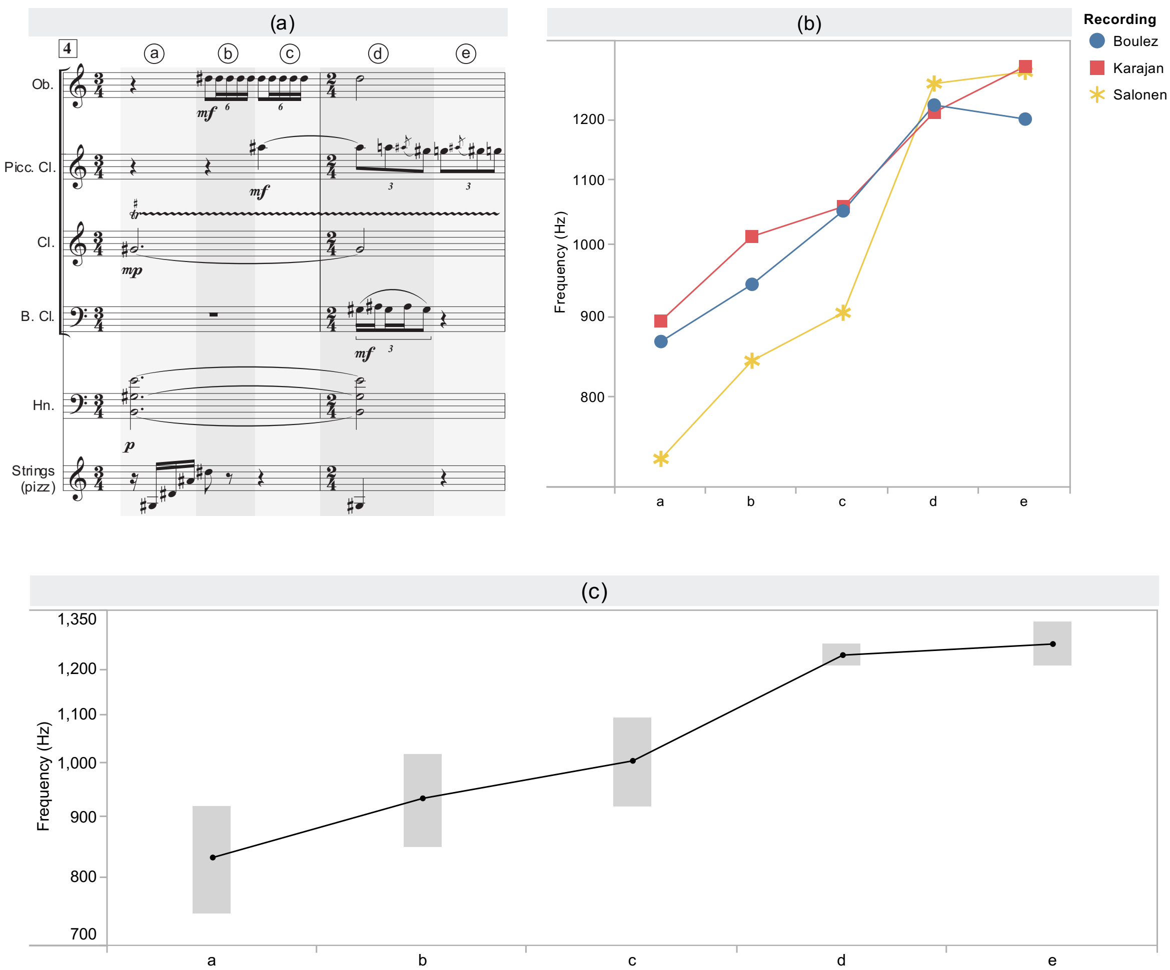


Figure 1-7. (a) Beat-level slices of the complete texture in Stravinsky's *Rite of Spring* (Intro to Part I, R4); (b) Centroid of each slice in three recordings; (c) Centroid of each slice averaged across recordings (+/- SD).

---

[14] Berlin Philharmonic & Herbert von Karajan, 1964 (DG 477 7160); Cleveland Orchestra & Pierre Boulez, 1992 (DG 435 769-2); Los Angeles Philharmonic & Esa-Pekka Salonen, 2006 (DG 483 9953).

The three recordings plotted in Figure 1-7b show considerable agreement, each enacting a process of timbral brightening over the course of the five beats. Discrepancies can be observed: Salonen's segments "a" through "c," for example, are notably darker than the other two recordings, and the ratio between "c" and "d" is much larger (SCQ=1.40) than the other two (avg. SCQ=1.16). As the texture-level centroid refers to the overall audio signal, many factors can explain these discrepancies. A viable approach, one which I take in Chapter 4, is to listen closely to the plotted audio snippets and consider potential sources of the variance—subtle differences in dynamic balance, usage of mutes, and recording conditions (e.g., mic placement, room size), to name a few examples, all play a part in shaping a recording's brightness in time.

An alternative approach to visualizing the texture-level centroid emphasizes agreement over variance: in Figure 1-7c I plot a single centroid trajectory corresponding to the average of the three consulted recordings. Grey vertical bars represent the standard deviation of the three recordings, with longer bars indicating a broader spread of possible realizations for that segment. In this visualization strategy, differences between recordings are less precisely defined, but the overall trend is more apparent.

The temporal scale of Figure 1-7 is that of the *beat*—one texture-level centroid per quarter note—but centroids can encompass much larger chunks of music. In Figure 1-8, I zoom out to analyze the relative brightness of three adjacent sections of *The Rite of Spring*. The two-measure excerpt previously discussed is part of a larger five-measure section (i.e., comprising the entirety of R4), which can be represented by a texture-level centroid and thereby compared to the formal chunks on either side of it. Leading up to R4 is a six-measure section, starting one measure after R3 ("Più mosso"), featuring English horn and bassoon trio. Following R4 is a three-measure section (R5) with chaotic flutes and clarinets supporting an oboe melody. Figure 1-8a presents the beginning of these three adjacent sections in reduced score. Figure 1-8b plots the texture-level centroid for each of the three sections, with the top panel broken down by recording and the bottom panel showing the average across recordings. The large-scale texture-level reading reveals a process of gradual brightening: each new section is on average brighter than the passage that precedes it, with a notable uptick at the onset of R4.[15]

These section-scale observations (Figure 1-8) demonstrate the system's potential to align brightness data with readings of form. I explore this idea in Chapter 4, wherein texture-level centroid readings are averaged over windows of time derived from a preliminary formal analysis of Webern's op. 6 no. 2. Connections between brightness and form are explored further in Chapter 5, wherein I integrate multiple textural levels in an analysis of the first movement of Varèse's *Octandre.*

[15] This observation may provide timbral support for narrative interpretations of this passage; in Straus' (2022) video analysis of the *Rite,* for example, he notes that in the A section (up to R4), "the world is still asleep in the predawn darkness," but at the onset of R4, "the darkness starts to break, and the world comes alive" (20:55).

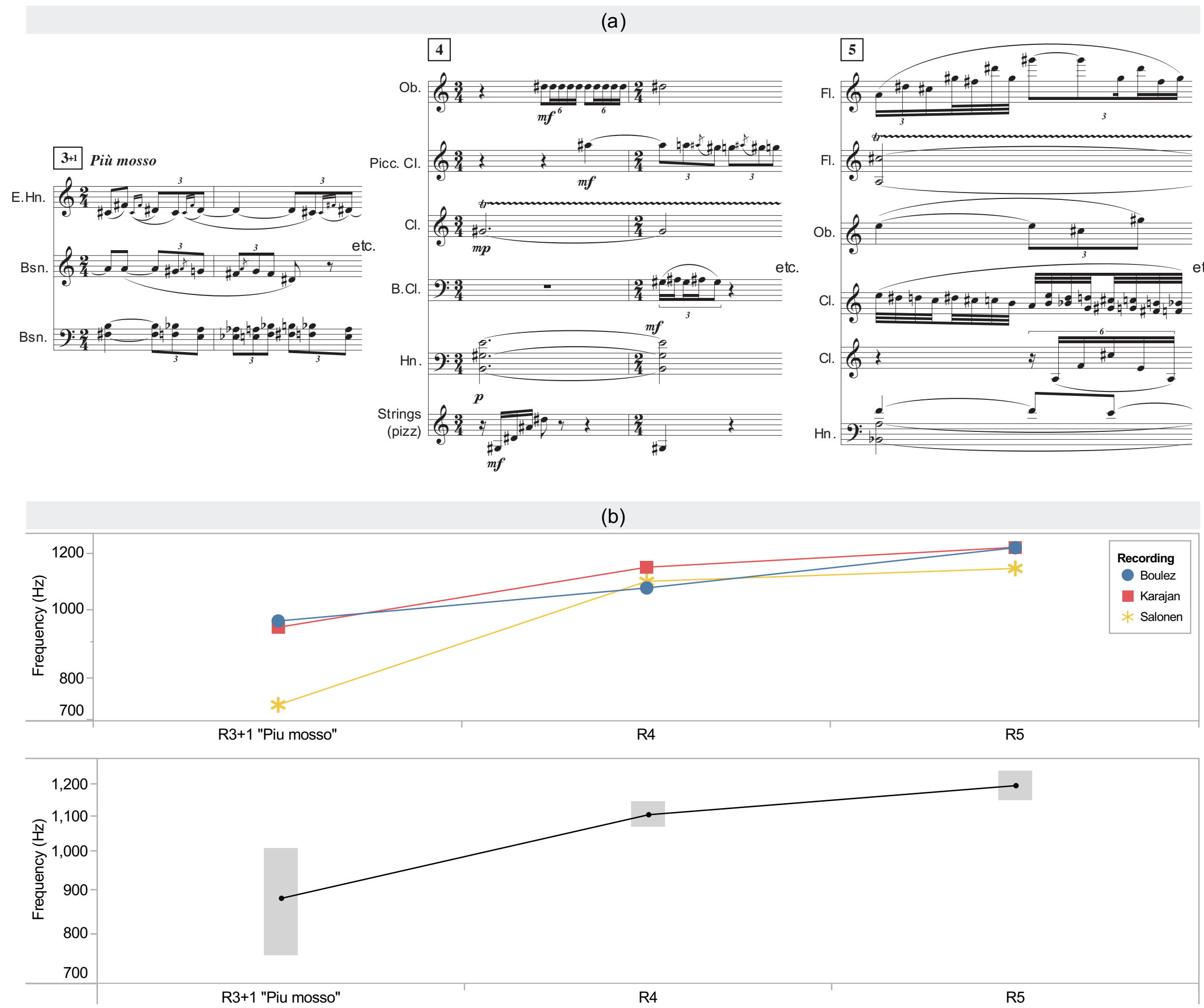


Figure 1-8. (a) Reduced score incipits of three adjacent sections in the Intro to Part I of Stravinsky's Rite of Spring (R3–5); (b) Texture-level centroid of those sections in three recordings; (c) Texture-level centroid of those sections averaged across recordings (+/- SD).

In the foregoing discussion, brightness analysis at different textural levels and temporal scales reveals very different timbral insights about the same passage of music. This points to the overarching benefit of the approach: the three textural levels—*element*, *layer*, and *texture*—are easily integrated and alternated with each other to make sense of multiple aspects of timbral design. Where one level falls short in explaining the perceptual effect of a passage, one or both of the other levels will compensate, thereby enriching our understanding of the music. For example, while the layer-level analysis in Figure 1-6 might explain well the relative brightness of polyphonic strands, the timbral makeup and internal contrasts of each of those strands will be better understood at the level of the element (e.g., Figure 1-4, 1-5). While texture-level analysis in Figure 1-7 quantifies "overall" brightness, the role of individual timbres and their combinations in bringing about that brightness can be elucidated at the lower levels.[16] There is thus an inherent flexibility in the approach, encouraging analysts to respond dynamically to the unique demands of the particular pieces analyzed.

Conclusion

The overall approach described above can be summarized as follows: through score-based analysis, I isolate discrete musical units at three textural levels—*element, layer,* and *texture*—and multiple temporal scales (narrow slices, larger segments, complete sections, etc.). I use orchestral sample libraries (for *elements* and *layers*) and recordings (for *textures*) to create audio files corresponding to these musical units. I measure the spectral centroid of the audio files to assess the brightness of individual units and calculate SCQs to quantify relative brightness between pairs of units. I freely alternate and integrate

[16] Occasionally, the different textural levels are essentially one and the same. In a monophonic passage for solo instrument, for example, there is there is only one "level," and it is simultaneously an element, a layer, and the whole texture. Similarly, in polyphonic music for solo instruments, layers are elements.

analysis at these different textural levels and temporal scales, attempting to capture the multiple ways that brightness shapes the music we hear.

This methodology is elaborated further and put to analytical work in the chapters that follow. In Chapters 2–4, each textural level will serve as the primary focus of an in-depth analysis, while Chapter 5 integrates all levels in a comprehensive brightness analysis. The next chapter homes in on element-level centroids in Arnold Schoenberg's "Farben," revealing patterns of concurrent and sequential brightness that go undetected by previous timbre-oriented approaches to the piece.

Chapter 2

Chord Color as Brightness in Arnold Schoenberg's "Farben"

This chapter analyzes the third piece of Schoenberg's *Five Pieces for Orchestra,* op. 16 ("Farben"). Unlike previous studies dealing solely with orchestration or comprehensive spectral analysis (Burkhart, 1973; Zeller, 2023; Cogan & Escot, 1976; MacCallum & Einbond, 2008), my approach focuses on element-level centroids, exploring the relative brightness of individual chord members. The analysis revolves around the idea that a chord's "color" is related to the distribution of its constituent element-level spectral centroids. Building on this, I develop methods for following chord-colors and individual voices through time, and briefly consider layer-level relationships between background and foreground material.

Existing Approaches to "Farben"

Often celebrated as the quintessential example of timbre-based composition, Schoenberg's "Farben" features a five-note chord subjected to almost continuous changes in orchestration. The substantial existing literature on the timbral dimension of the work seems either to provide too much or too little detail. In this section, I review two prevalent analytical approaches to timbre in "Farben": *instrument-oriented* approaches (Burkhart, 1973; Zeller, 2023) and *spectrographic* approaches (Cogan & Escot, 1976; MacCallum & Einbond, 2008). I propose an alternative approach that employs concurrent element-level centroids, allowing us to examine pertinent acoustic details without being overwhelmed by spectral information.

Nearly all preexisting timbral analyses of “Farben” orient themselves toward Schoenberg’s selection and combination of musical instruments, homing in on the changing instrumental makeup of the nucleic five-note chord. The resulting auditory effect of the instrumentation, however, is either discussed only in cursory terms, or left to be extrapolated by the reader. Burkhart’s (1973) comprehensive analysis of the piece is a prime example of this approach. In his analysis of the instrumentation (pp. 151–166), Burkhart lists meticulously every instrument contained in each instance of the five-note chord (which he calls the “organism,” a term I adopt in this chapter) and reveals a partially serial ordering of instruments at the precise midpoint of the piece. Figure 2-1 reproduces a relevant segment of Burkhart’s multipage diagram, zooming in on the first leg of the “partially serial” arrangement of instruments in the upper quartet of the five-note chord. The sequence of vertical slices numbered 1–11, which Burkhart calls the “prime series,” is repeated in retrograde (with 6 missing, and with 4 and 5 switched). After this imperfect retrograde, the prime series is then repeated in full and once more in part (not shown).

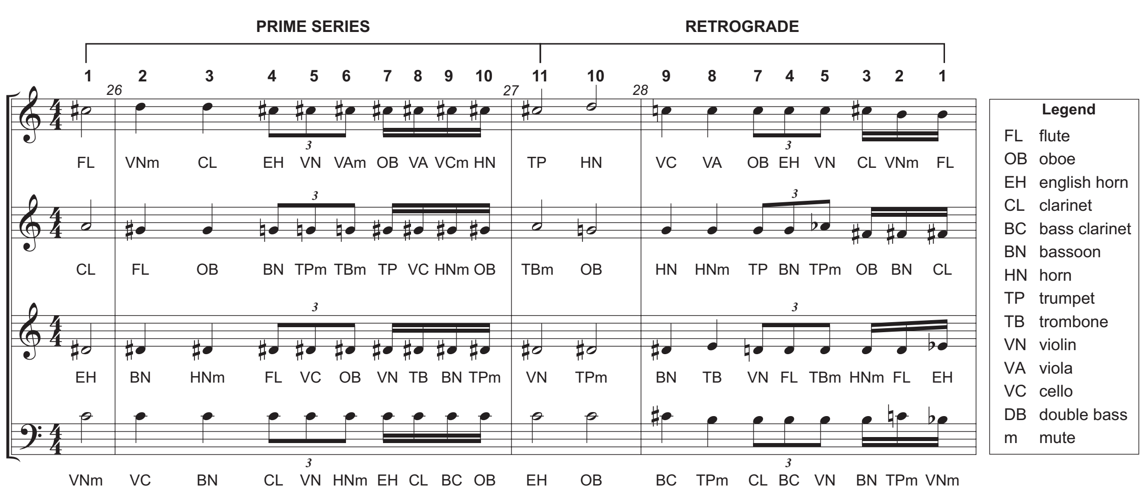


Figure 2-1. Engraved segment of Charles Burkhart’s Fig. 9 (1973, p. 159), top four staves only. The series of instrumental combinations numbered 1–11 is repeated in retrograde.

In his discussion of timbre in the piece, Burkhart makes an important observation which in my view is not led to its correct conclusion. He states: "…fifteen instruments … participate in the organism. But it is quite wrong to assume only fifteen colors" (p. 154). He elaborates that the use of mutes and special techniques (e.g., harmonics, pizzicato) constitute distinct "colors" or, more accurately, "color components." He goes on to argue convincingly that, despite this, Schoenberg likely took instrument types (regardless of mutes, etc.) to be the "color components of highest status," relegating timbral modifiers like mutes to a secondary role. Indeed, for Burkhart's serial ordering shown in Figure 2-1 to work, we must consider the horn and muted horn, for example, to be the same entity. However, the initial observation—that instrument type is not synonymous with timbre—ought to be extended beyond the use of mutes and special techniques: indeed, all pitched musical instruments can produce a wide range of unique timbral qualities that covary with pitch and playing intensity. Siedenburg and McAdams (2017) offer a useful biological metaphor: "a single type of sound-producing object … may give rise to a timbral *genus* that encompasses various timbral *species*" (p. 3).[1] Taking this into account, Burkhart's analysis can be said to uncover meaningful structures at the broad level of the timbral *genus* (perhaps intended by Schoenberg). But there may be significant "interspecific" patterns at play—patterns that become visible when one considers timbral categories that cut across instrument types.

[1] Sandell and Chronopoulos (1996) employ the term *macrotimbre* to describe the complex of timbres that can be produced by a single sound source. Soden (2020, pp. 116–124) uses the term *metatimbre* for groupings of timbres characterized by shared attributes (including, but not limited to, those produced by the same instrument).

More recently, Zeller's (2023) analysis of "Farben" moves beyond the changing colors of the organism, shifting focus to the textural and timbral details of the often-neglected foreground material of the work. Zeller's novel approach defines a lexicon of timbral functions that have analogues in the realm of pitch. Central to his discussion of "Farben" is a cluster of concepts he calls *timbrality*, *timbralitization*, and *timbral modulation*—analogues to tonality, tonicization, and modulation, respectively. However, rather like Burkhart, timbre is not quantified but rather indexed through an account of the work's instrumentation. For example, a major moment in Zeller's analysis is the change that occurs at m. 11. The claim is that the timbrality (i.e., the prominent referential timbre that characterizes the passage), which was previously located in the woodwind-dominated five-note chord, is temporarily shifted to the string instruments and a timbral "cadence" is reached. Such a claim relies heavily on the generic characteristics and behaviors of specific instrument families (e.g., winds vs. strings) to define and describe timbre, as opposed to quantifying the acoustic and perceptual effect of such characteristics and behaviors. This method thus keeps acoustic details at arm's length, obscuring the potential nuances and surprising details that an acoustically informed analysis can reveal.

A much smaller collection of published work analyzes "Farben" from an acoustic standpoint. In most cases, this work ends up being overly inclusive, producing for example the spectrograms of each instrument in every chord, along with the composite spectrum of the blended result. Cogan and Escot's (1976) discussion of the opening two chords of the piece (pp. 365–368) is one of the best-known analyses of this type. The authors use acoustic data from Seashore (1938) (supplemented by Ancell, 1960), to plot the spectral energy distribution of each instrumental tone in both chords. The data, laid out in columns for each instrument, gives the percentage of the total sound concentrated in each overtone, with an extra column representing the composite spectrum of the blended result (i.e., the chord). These complicated graphs yield broad conclusions: while the first sonority has a narrower overall span, and its energy is concentrated in the fundamental frequencies (i.e., in the low- to mid-range), the second sonority has a wider span, with energy centered around the second and third partials (i.e., in a higher register than the first sonority). While their approach offers detailed insights into the spectral energy distribution of the opening chords, it highlights the tradeoff between legibility and exhaustiveness—a recurring challenge in timbral analysis.

To be sure, a major advantage of the acoustic approach is its ability to analyze features of the composite result (i.e., the total spectrum of each "Farben" chord). For example, Cogan and Escot use spectral information to describe the prominent beating that occurs in certain regions of the spectrum (p. 373). MacCallum & Einbond (2008), similarly, overlay the complete spectra of the two chords on the same spectrograph, allowing for visual judgment of the different spectral peaks in each.[2] Siedenburg (2008) visualizes the timbral difference between the first two chords by combining five audio features onto a single

[2] This visualization is not the centerpiece of MacCallum & Einbond's analysis but is rather used to concretize a somewhat more easy-to-digest claim, namely that the second sonority is perceptually "rougher" than the first one.

graph (spectral centroid, bark-flux, flux, skewness, and roll-off), sacrificing readability for increased information density.

In their full analysis of the piece (pp. 412–426), Cogan and Escot devise a brightness system that on the surface resembles the one presented in this chapter, albeit with some crucial differences. The authors devise a “timbral scale” by arranging the spectrum of each instrument from darkest to brightest. Then, the five-note chord is followed through its changing orchestrations. There are two major problems with this approach. The first is that the spectra are taken from a single pitch (G4) and used to make a general characterization about the overall brightness of each instrument as a whole. This ignores the diverse range of timbres (i.e., timbral “species”) producible on a single instrument; all string instruments, for example, are inaccurately considered “bright” whenever they appear, regardless of pitch. The second is that these judgments about overall brightness appear to have been made optically, and not computationally (i.e., relying on visual estimation of spectrographic data rather than using the spectral centroid or another statistical measure). The resulting analysis is thus an instrument-oriented approach in an acoustic guise, in that the analysis amounts to a list of which instruments are present in each chord, with the added information about the brightness of the overall macrotimbre of the instrument.

To summarize: in instrument-oriented approaches, too much is left to the reader, who must triangulate timbre using given information about instrument, pitch, and dynamic. In spectrographic approaches, on the other hand, timbral analysis is grounded in spectral content, but perceptually and analytically relevant features are buried in the details.

Responding to this trade-off, my approach to the piece focuses on the relative brightness of elements, remaining rooted in empirical measurement without presenting an excessive amount of acoustic data.

The analysis that follows thus begins with a close reading of relative brightness in the opening chords of the piece. Then, to encapsulate the changing character of the “Farben” chords over time, I develop a framework I call “chord color space,” based on the average element-level brightness of each chord, along with two aspects of element-level centroid distribution which I call *span* and *unevenness*. I complement this synoptic approach with a more zoomed-in one, tracing the development of each individual contrapuntal line through time using sequential element-level centroids. Finally, I consider layer-level centroids of the chords and examine relationships between the background and foreground elements of the piece.

Chord Colors and their Composition

Figure 2-2 plots the spectral centroid of each instrumental tone in the opening chords of the piece. Colors on the graph match spectral centroids to their respective fundamental frequencies, shown in music notation below. This kind of visualization, which I call a *concurrent element graph,* allows us to make detailed analytical claims about how element centroids affect both the character of each chord and the degree of the contrast between chords. These details will be explored below.

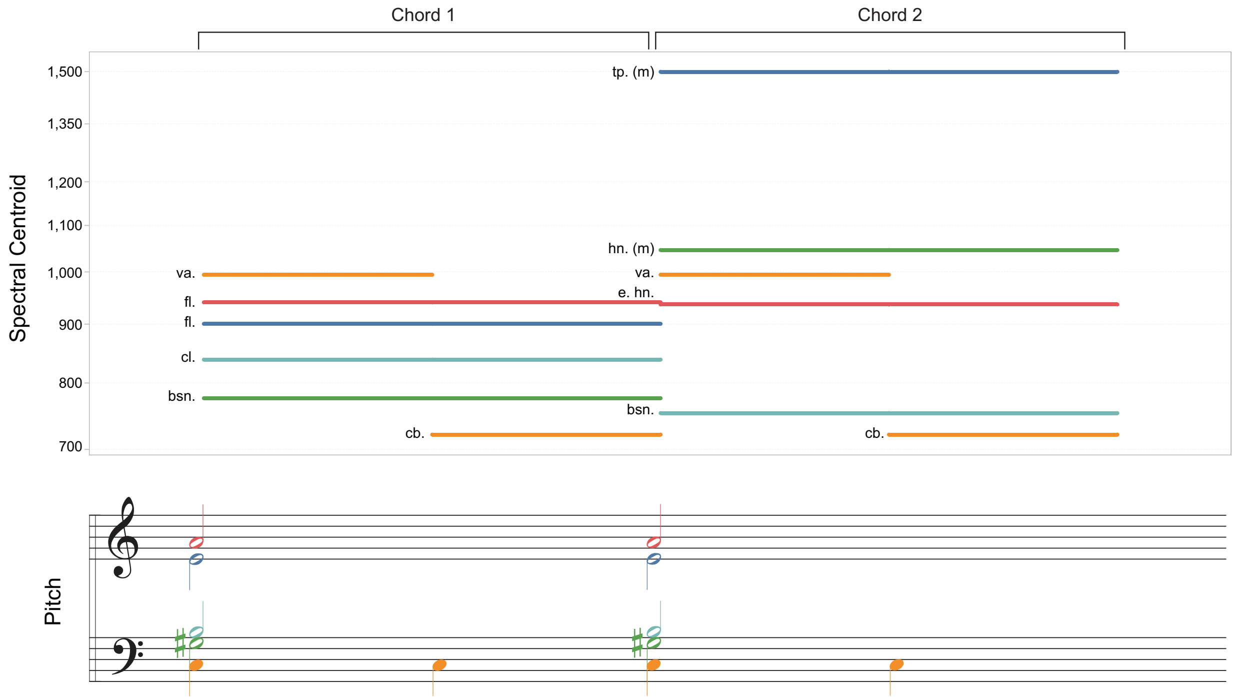


Figure 2-2. Concurrent element graph of the opening two sonorities in "Farben." The first chord is evenly spaced, while the second is highly clustered. Instrument samples from Garritan Personal Orchestra 5.

Table 2-1. (a) Unevenness scores for Chords 1 and 2. (b) Silhouette scores of k-means clustering. Element-level centroids in Chord 1 do not appear to be strongly clustered, while elements in Chord 2 appear to form three clusters.

	(a) Unevenness Scores	(b) Silhouette Scores of K-means Clustering			
		K = 2	K = 3	K = 4	K = 5
Chord 1	0.014	0.53	0.28	0.19	0.10
Chord 2	**0.18**	0.48	**0.66**	0.39	0.27

Figure 2-2 allows us to observe readily both the overall span and distribution of the element-level centroids that make up each chord. The opening sonority, for example, appears to occupy a narrow range, with rather evenly spaced component centroids; the second chord spans a wider range and has an uneven distribution, forming clusters of elements. The overall range of the element-level centroids—an estimate of how much spectral space the sonority takes up—can be measured by examining the SCQ that obtains between the lowest and highest centroid. In this case, the first sonority spans a low SCQ (1.38) and the second a high SCQ (2.08). In addition, our visual estimate of the distribution can be confirmed by analyzing the standard deviation of adjacent SCQs in each sonority. By "adjacent" I mean only the SCQs that obtain between neighboring centroids in the stack. A standard deviation of zero would mean that all adjacent SCQs are of the exact same size, indicating a maximally even spacing of centroids.[3] I call this measure *unevenness*.

Table 2-1a shows unevenness scores for the first two sonorities. In line with our visual estimate, the unevenness of the first chord is very low (0.014) when compared to the second chord (0.18).[4] I will proceed under the assumption that when centroids of individual elements are spread out and unevenly distributed, as they are in the second chord, the overall sonority becomes less perceptually uniform than it would be if the same number of centroids were tightly packed and evenly distributed.[5] One way of understanding the

---

[3] The possible range of adjacent SCQs is small, so the standard deviation is commonly lower than 1.

[4] For these calculations, I am treating the "Farben" chords as six-note sonorities, by including both instances of the bass note (C3) as distinct elements, even though these two sounds are presented successively.

[5] This aligns only loosely with empirical literature (Sandell, 1995), which demonstrates that the degree of blend correlates with proximity of centroids: in this case, while overlapping or adjacent elements in an unevenly spaced sonority may indeed blend, the overall sonority will exhibit higher centroid ratios (SCQs) between those clusters, resulting in elements that remain distinct and unblended. While additional empirical evidence is needed here, my hypothesis is that sonorities with low unevenness scores will result in a more homogeneous overall sound and vice versa.

opening section of the piece, then, is that a perceived shimmering of the opening chords is partially achieved via a regular alternation between homogeneity (low span and unevenness) and heterogeneity (high span and unevenness).

High unevenness typically indicates the presence of centroid clusters, but the *unevenness* value itself does not provide information about the number or size of such clusters. This would seem to be a crucial consideration, both perceptually and analytically: if the voices in these sonorities do not cohere into a single unit, how many distinct elements can be discerned? In my approach to this piece, a sonority's cluster count is both visually estimated by examining a concurrent element graph and further corroborated computationally using *k*-means clustering and silhouette scores. In Figure 2-2, we can see that the second sonority appears to visually divide into three clusters: (1) muted trumpet; (2) muted horn, viola, and English horn; and (3) bassoon and double bass. Performing *k*-means clustering on the element-level centroids of the chord with a *k* of three supports this observation.

We can confirm that three is a particularly optimal number of clusters for this data by examining the silhouette scores of computational clustering solutions with different *k* values. [6] Table 2-1b shows the silhouette scores for each clustering solution: none of the

---

[6] *k*-means clustering iteratively assigns data points to clusters based on their distance from a central point in each cluster. The parameter "k" denotes the desired number of clusters, chosen in advance by the analyst. Silhouette scores assess the cohesion and separation of the clustering solution for each "k" value; a high silhouette score (closer to 1) suggests that the tones in each cluster are closer to each other and further from those in other clusters.

silhouette scores for Chord 1 exceeds 0.53, indicating that no clustering solution is particularly optimal. Chord 2, by contrast, gets a higher silhouette score of 0.66 for a three-cluster solution, suggesting that it can more plausibly be understood as clumping into three. Clustering analysis of this sort is useful for determining the perceptual character of a sonority, particularly as regards the single instruments or instrument groups that might stand out from the crowd. For example, it stands to reason that the muted trumpet's centroid, forming its own separate "cluster" in the second sonority, will be less securely tethered to the rest of the ensemble, and thus will stand out from the sonority. The bassoon, too, isolated for the first beat of the measure, can be understood as an isolated and distinct timbre until it is joined by the double bass on the second beat.

This approach to a sonority's clustering, where outliers can be understood as perceptual standouts, is pertinent to contemporary orchestration literature focusing on perception. McAdams et al (2022) introduce a three-prong taxonomy to describe the perceptual grouping of simultaneous sounds: *augmentation*, wherein one instrument dominates and is colored or embellished by others; *emergence*, where the constituent elements of the blend are not identifiable; and *heterogeneity*, where no perceptual integration occurs. The authors clarify, "One can have various degrees or strengths of blending along a continuum from completely fused to completely segregated" (4.3). The clustering approach described above can allow for nuanced analysis of a whole range of those intermediate, somewhat

---

while a low silhouette score (closer to 0) indicates that the clusters may overlap. See Rousseeuw (1987).

heterogenous combinations that fall between category boundaries, providing insight as to which instruments stand out and why.

Also of interest, and particularly visible in Figure 2-2 is the ordering of the five contrapuntal voices in the organism. One might expect the ordering (e.g., from high to low) of centroids to be the same as the ordering of the voices (e.g., with the soprano's centroid on top, the bass on bottom, etc.), but that appears to be the case only infrequently in the piece. In the first chord, for example, the centroid analysis gives the surprising result that the lowest sounding pitch, C3, is simultaneously the most bright and dark element, being played first by the solo viola, and second by the double bass.[7] Harmonically, C3 is the bass voice, but from a brightness perspective, C3 is first the "highest" (brightest) and then the "lowest" (darkest) sounding note in the chord. The ordering of voices is considerably dissimilar in the next sonority. One way of understanding the transformation of voice order is to point out that the alto voice (E4) and the lower tenor voice (G♯3) both become considerably brighter: the alto voice traverses a large SCQ (1.66), jumping to the top of the sonority, while the lower tenor voice brightens by an SCQ of 1.34, now sitting above the viola's C3. At the same time, the higher tenor voice (B3) is darkened slightly. The movement of each voice and the resulting order is summarized in Figure 2-3. Solid lines indicate a voice that occupies the same brightness region in both sonorities, while dotted lines indicate a more noticeable

---

[7] This observation is directly conflicting with Cogan & Escot's analysis of the piece, which makes no distinction between these two string instruments, grouping them together into a single step on their "timbral scale."

change in brightness. This approach will be developed later in this chapter, when we delve into the *sequential* centroids that make up each polyphonic voice's trajectory through the changing organism.

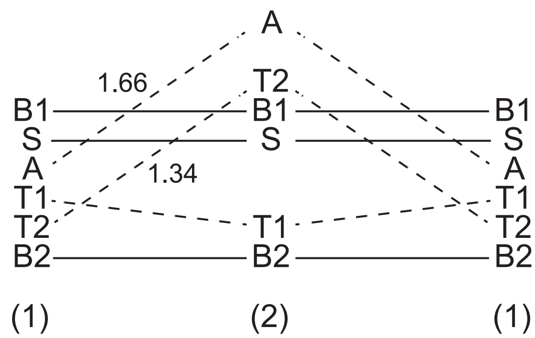


Figure 2-3. Changing ordering of the voices in the "Farben" chords (labeled 1 and 2), showing the alto voice and low tenor voice shifting to the top of the sonority.

Figure 2-4 synthesizes the foregoing observations regarding the element-level centroids of the opening two sonorities. The figure replicates the concurrent element graph from Figure 2-2 but incorporates additional analytical overlay regarding the span and unevenness of the sonorities. Colors reveal the number of clusters in each sonority. Voice labels S, A, T1, T2, B1, and B2 match the centroids to their respective pitches shown in music notation below. "Stand out" elements—those instruments with centroids conspicuously distant from the rest of the ensemble—are encased in dotted lines.

The observations above—that the opening sonorities differ in terms of their homogeneity, spacing, cluster-count, distinctness of individual elements, and voice order—point to the kinds of structural details we might miss when we fail to attend to and quantify the timbres that comprise "Farben." The approach, although it focuses on just the opening sonorities, illuminates how subtle orchestration choices significantly impact the character, or "color," of the “Farben” chords in ways that cannot readily be teased out via instrument-oriented score summary or comprehensive spectrographic approaches. In addition, it points to ways of analyzing how these colors undergo transformations and thereby give shape to the piece in time. This last point will be developed further in the next section below.

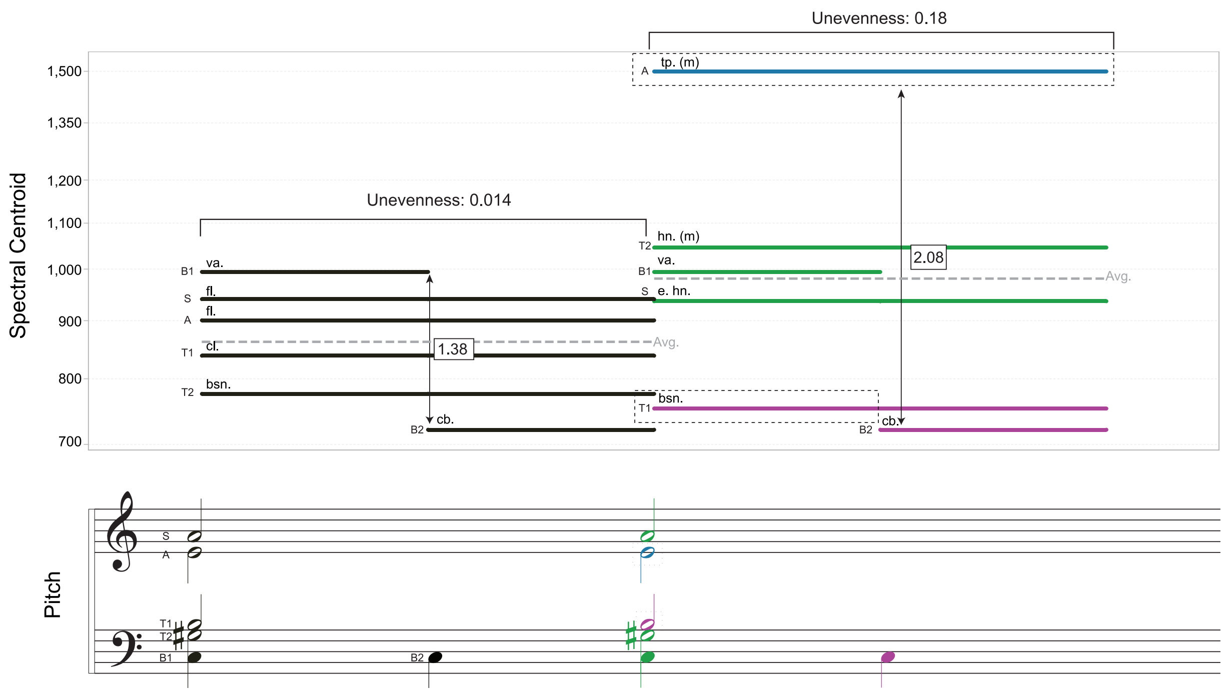


Figure 2-4. Annotated concurrent layer graph showing unevenness, span, voice ordering, clustering, and stand-out timbres. The first sonority is even, unclustered, and spans an SCQ of 1.38; the second sonority is uneven, clustered into three, and spans a SCQ of 2.08.

It should be noted that element-level centroids in a blended chord may not be “directly” perceived. That is, one does not necessarily hear the “Farben” chords as composed of multiple distinct brightness values, in much the same way that one does not directly perceive the color green to be a combination of blue and yellow in some proportion. Rather, the *effect* of the combination is perceived. The overall timbre of a chord—e.g., the global brightness value—*supervenes* on the timbres of its elements, which is to say that a change in global brightness can only occur if there is a change in brightness at the element level. Thus, it can be revealing to analyze perceived timbral wholes by studying their "unperceived" timbral parts. In much of the analysis that follows I will attempt to thread the needle between that which can and can't be directly perceived in the "Farben" chords through time: building on my analysis of the first two chords above, I examine minute variations of element-level centroids in the chords as they change through time. I then coordinate these lower-level variations with the more directly "perceptible" brightness value captured by the spectral centroid of the full chords, and thereby discuss interactions between textural layers.

Colors in Time

In this section I analyze changing chord colors through larger chunks of music, and in so doing point to a few possible ways forward for the analysis of timbre and musical form that will be taken up again in Chapter 5. As I showed above, my understanding of the "color" of a chord in "Farben" will depend in large part on the relationships that obtain between its element-level centroids: their span, unevenness, cluster count, and voice ordering. I will begin by analyzing how these features combine in a major timbral contrast the occurs in the first section of the piece, before tracing their progression through a longer passage (mm. 13–23).

The well-known first section of the piece features a canon wherein each voice in the opening five-note sonority takes a turn carrying out a simple two-part motion that displaces it a semitone lower: up a semitone, then down a whole tone. With each voice performing its iteration of this motive, the "Farben" chord crawls slowly downward, eventually finding itself a full semitone lower (T_{11}) by the ninth measure of the piece.[8] As this canonic process unfolds, the chord oscillates back and forth ten times between the two instrument combinations described above. The section closes in measure eleven with what Burkhart calls a "string punctuation," and what Zeller calls a timbral cadence in the "timbrality" of the strings: the chord (T_{11}) is transposed down an octave and given to a choir of solo cellos, with a contrabassoon playing the bass note.

Beyond pointing out merely that this punctuating timbral shift comprises a change in instrumentation, we can build on the approach described above and find prominent changes among element-level centroids. To this end, Figure 2-5 summarizes the first section of the piece (mm. 1–11), reproducing the element-level centroid data of the first two chords, juxtaposed with the string sonority in measure eleven. To save space, sonorities are summarized here with single data points with their durations and repetitions removed. The repeat sign indicates the continued alternation of the first two chords. The most visually apparent feature of the graph is that the string sonority (labeled b) is on the whole much darker than either of the previous two sonorities (labeled a)—indeed, the mean element-level centroid sits around 400 Hz, much lower than that of both the first and second sonority (863 Hz and 993 Hz, respectively).[9] The downward leap in pitch thus

---

[8] For an overview of the large-scale pitch content of the work, tracing this canon through the piece, see Rahn (1980, Ch. 4). See also Burkhart (1973).

[9] The string samples I consulted here are the "Solo Cello 1" samples in Garritan Personal Orchestra 5. Other cello samples in this very library emphasize the noisy brightness of the

occurs in tandem with a drastic timbral darkening of each element in the sonority. The graph also makes the homogeneity of the string sonority explicit; its span is narrow, and its unevenness is somewhat low (0.14).[10]

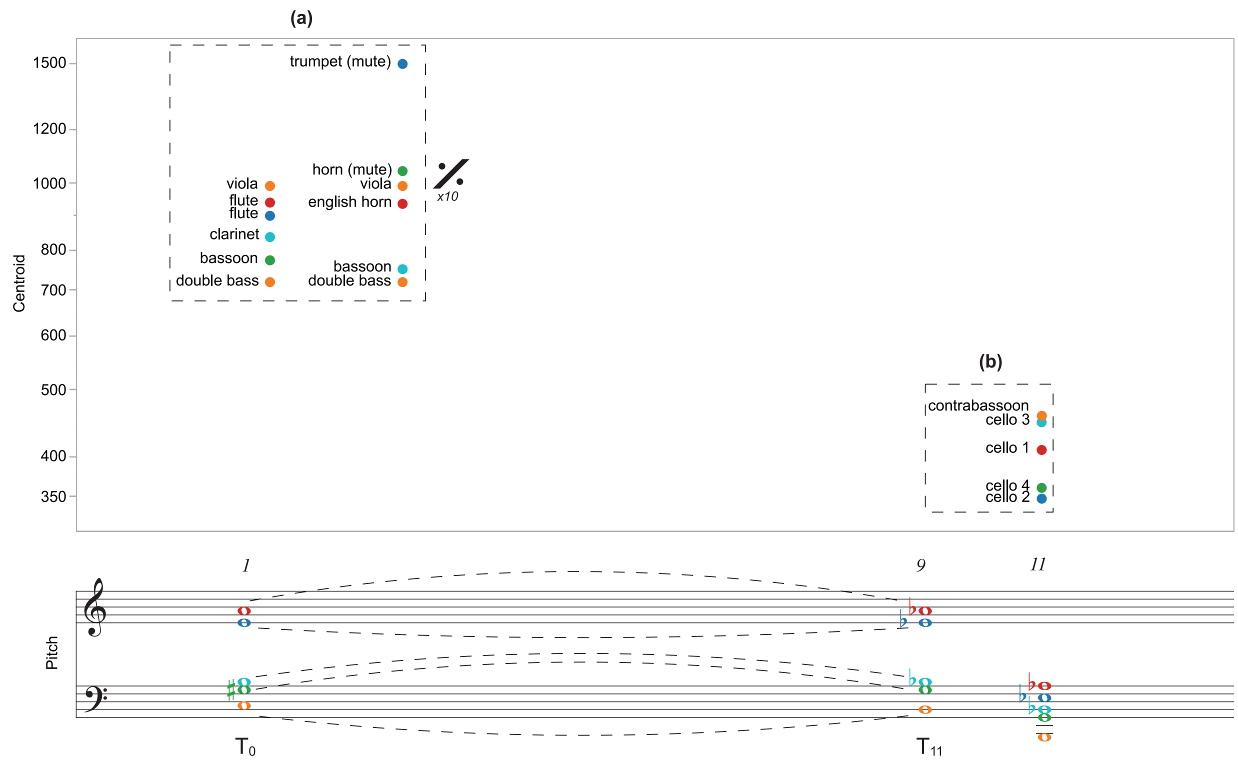


Figure 2-5. Concurrent element graph showing color differences between the repeating opening sonorities and the string punctuation in measure 11. The string punctuation's average element-level centroid is much lower, and its span narrower.

---

bow sound, yielding different analytical results. Schoenberg calls for solo instruments, one to a part. The "Solo Cello 1" samples thus seem appropriate, as they are intended to be used as solo instruments in chamber music settings, rather than, say, the "Cello Player 1" and "Cello Player 2" samples which are intended to be combined with each another in unison to form a large string section.

[10] In this case, with a span so low, it seems to me less pertinent to discuss voice order and cluster count, given that minute changes would bring about different clustering results.

Three factors described above—average element-level brightness, span, and unevenness—can be combined in a single visualization scheme that I call "chord color space" (Figure 2-6). Points in this space represent sonorities. The position of the point along the y-axis represents the average element-level centroid of the sonority, while the x-axis denotes the overall span. The shading of points represents unevenness (with lighter points indicating greater unevenness). Plotting the three sonorities in this space (numbered 1–3) illuminates the degree to which the third sonority is dissimilar to the first two, and the element-level factors that contribute to that dissimilarity.

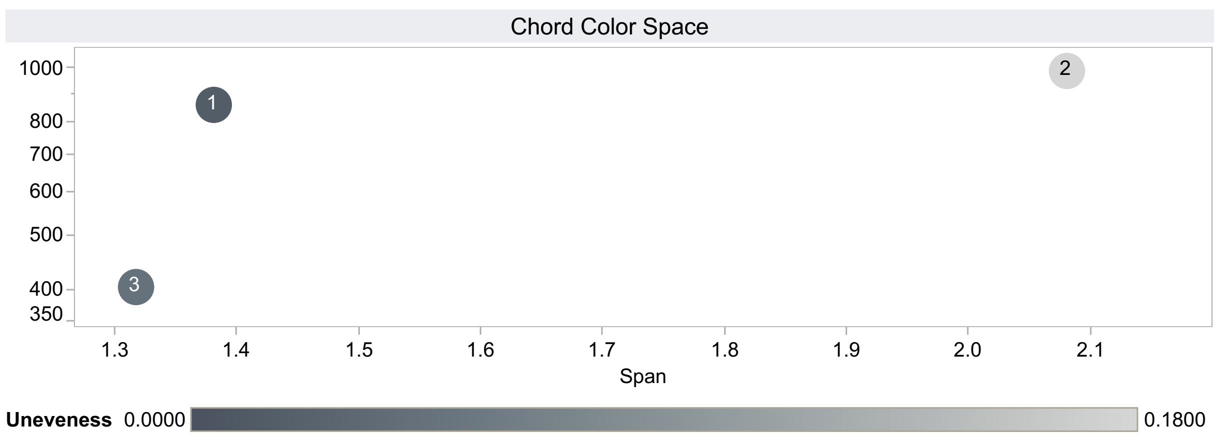


Figure 2-6. The first three instrument combinations in "chord color space," showing average element-level centroid plotted against span. Lighter shading indicates higher unevenness values. The third combination (3) has a significantly lower average element-level centroid compared to the first two combinations (1 and 2). Combination 2 has a much higher span than the others.

I want to explore further the idea that a succession of orchestrated chords might be reframed as a path through chord color space. To that end, let us examine the orchestrations of the "Farben" organism that initiates the second section of the piece, immediately following the string punctuation discussed above. In mm. 13–14, the semitone displaced iteration of the opening chord (T_{11}) is treated to four different orchestrations, as the bass voice moves up a minor third in m. 14, anticipating the wholesale upward transposition of the organism in m. 15. Figure 2-7 presents element-level centroid data for the four sonorities in mm. 13–14. The top panel (a) shows a concurrent element graph, and the bottom panel (b) plots the sonorities in color space.

Both panels of Figure 2-7 reveal a clear alternation between low-span and high-span chords, continuing the pattern established in the first measure of the piece: the two downbeat sonorities (i.e., 4 and 6) receive a more tightly packed orchestration than the sonorities that fall on the second half note of each measure (i.e., 5 and 7).[11] Within the fourth and seventh sonority, the bass voice enacts a darkening gesture, again following a precedent set in the first measures of the piece but with different instruments: in both of these sonorities (4 and 7) the bass voice begins with the muted horn and moves to the bass clarinet on the second beat.[12] Each sonority has one or more "stand-out" timbres, as indicated in Figure 2-7a with dotted rectangles. Over the course of the short passage, each of the four voices in the upper quartet gets a turn being prominent: the soprano voice (bassoon) in the fourth sonority, the lower tenor voice (horn) in the fifth, the alto voice (oboe) in the sixth, and the higher tenor voice (trumpet) in the seventh sonority.

---

[11] Note that the numbers I use to refer to sonorities in this chapter are offset from Burkhart's similar numbering system by one, because I count the string punctuation.

[12] This is part of the bass voice's repeating three-beat pattern that goes *muted horn* → *bass clarinet* → *bassoon*, out of phase with the upper quartet's two-beat timbral rhythm.

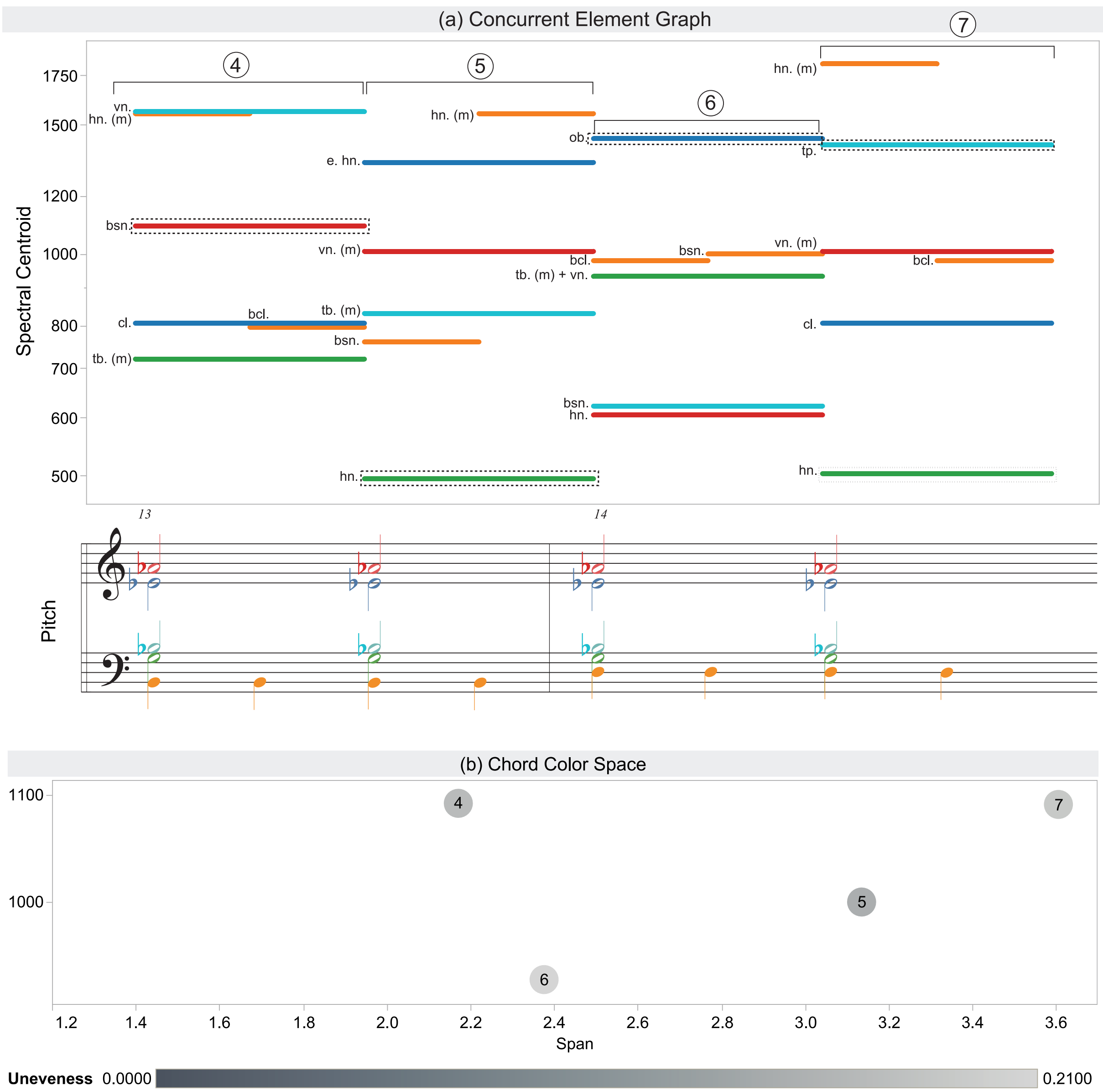


Figure 2-7. (a) Concurrent element graph of four sonorities in mm. 13–14. (b) The same four sonorities in chord color space.

All four sonorities in mm. 13–14 are highly clustered, with relatively high unevenness scores (represented by lighter shading in the bottom panel). Interestingly, the bass voice appears to be a key culprit in this unevenness: were we to isolate the top quartet, we would find much more evenly spaced element-level centroids (save for the sixth sonority). Thus, the bass voice—which Burkhart points out is a wholly separate musical entity than the upper quartet—acts as a kind of corruptor of the upper quartet in this passage, complicating the otherwise even spacing of its centroids and making the resulting sonority less unified.

In measure 15, a new transposition of the opening chord is reached (T_2). The chord continues to be reorchestrated every half note as its pitches remain motionless in mm. 15–16. A new instance of the canon begins in m. 17 but fails to fully materialize, ending up back at T_2 in m. 19 and remaining still for two more measures (before the true canon is carried out in half notes in mm. 21–23). The relative stasis of the pitches in this passage is countered by a remarkable amount of timbral activity, both within the organism and in the overall texture, as new musical layers are added. Figure 2-8 summarizes the effect of the orchestration changes in mm. 15–23 by plotting each sonority as a point in chord color space. I number the sonorities in this passage 8–24, continuing the numbering from previous figures.[13]

[13] The bass voice is not factored into this figure, as it appears to follow a separate logic, changing timbres on every quarter note against the upper quartet's half notes, and looping a three-quarter instrumental pattern.

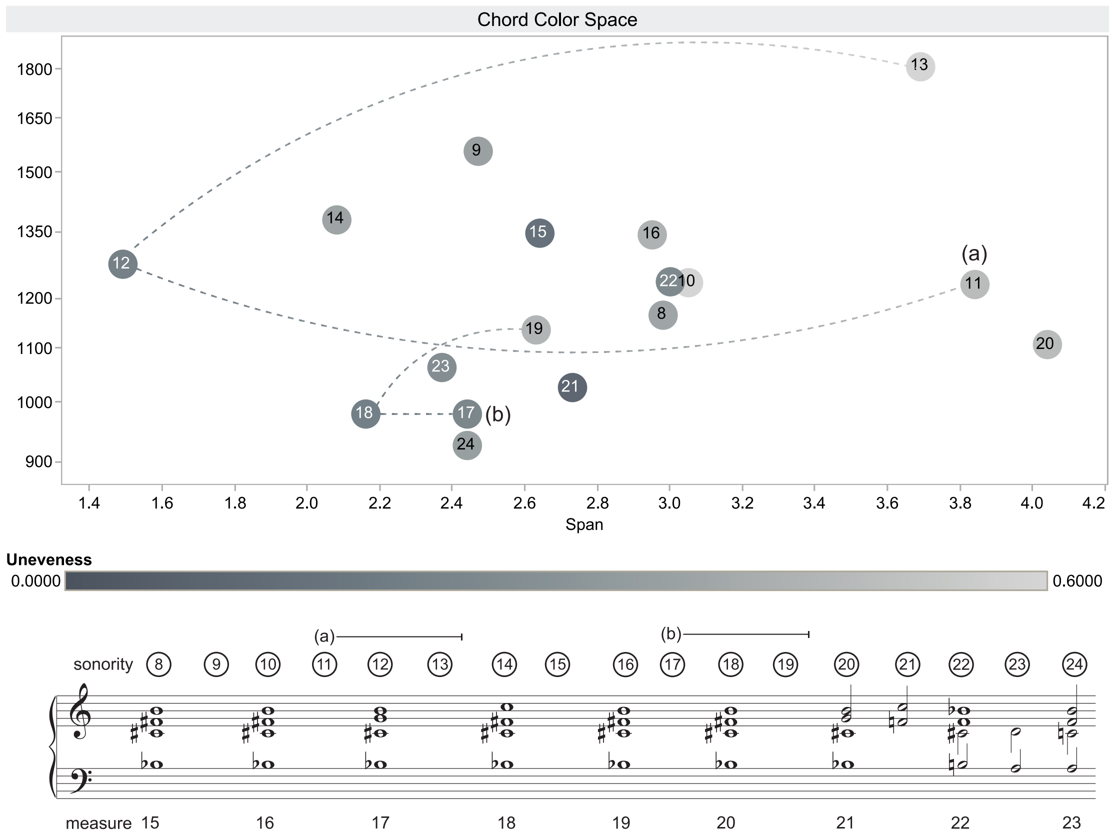


Figure 2-8. Sonorities 8–24 (mm. 15–23) in chord color space, with two opposing three-chord segments labeled (a and b).

The color variance from chord to chord is evident in Figure 2-8: adjacent sonorities are very rarely close to each other in all three dimensions of the space. Two extreme segments are worth singling out (labeled a and b). The path connecting sonorities 11, 12, and 13, beginning in the second half of m. 16 (labeled a), is the most dramatic in the passage: while the link between sonority 11 and 12 does not rise in average brightness, it carries out a drastic reduction in both unevenness and span. Then, to get from 12 to 13, these timbral changes are undone, and the sonority becomes much brighter overall. A very different kind of path is observed from 17–19 (second half of m. 19, labeled b), revealing three adjacent sonorities that are very closely related in terms of chord color. The two segments can each be viewed as a reversal of activity occurring in a foreground layer. In the first case (mm. 16–17), piccolo and celesta are added to the texture on the repeated pitch B5, floating motionless above the organism: while this new layer is static, the organism undergoes dramatic changes in response to it. The second case (mm. 19–20) is the opposite: a more complex and active foreground layer—the so-called “jumping trout” motive—is added to the texture, as the organism moves less dynamically. The brightness relationship between the organism and these added layers will be explored later in this chapter (Figure 2-12 below).

In m. 24, the organism shifts again to a new transposition of the opening chord (T_4). The music then begins to unravel: repetitions of the canon—at ever-increasing speeds—overlap with each other to bring us lower in pitch, as changes in orchestration occur more frequently (with every sixteenth note in m. 26 and mm. 28–29) and foreground interjections become more incessant. This dizzying passage lands at m. 30, when a string choir produces the opening chord in its original transposition (T_0), signaling the onset of the final section of the work.

In m. 32, the canon is initiated once again, but this time it is inverted: each voice moves *down* a semitone, then *up* a whole tone, thereby inching the organism upward. As we’ve seen, the norm in the first section of the piece is to alternate low-span and high-span sonorities, with the former tending to fall on downbeats—this was particularly apparent in the opening alternation (mm. 1–11) and in mm. 13–14 (Figure 2-7 above). Here at m. 32, in tandem with a motivic inversion, the established span pattern is “inverted,” with more high-span sonorities falling on downbeats, and low-span sonorities on upbeats. Figure 2-9 displays a graph of measure 32, a prominent instance of widely spaced centroids becoming narrower in the second half of the measure.[14]

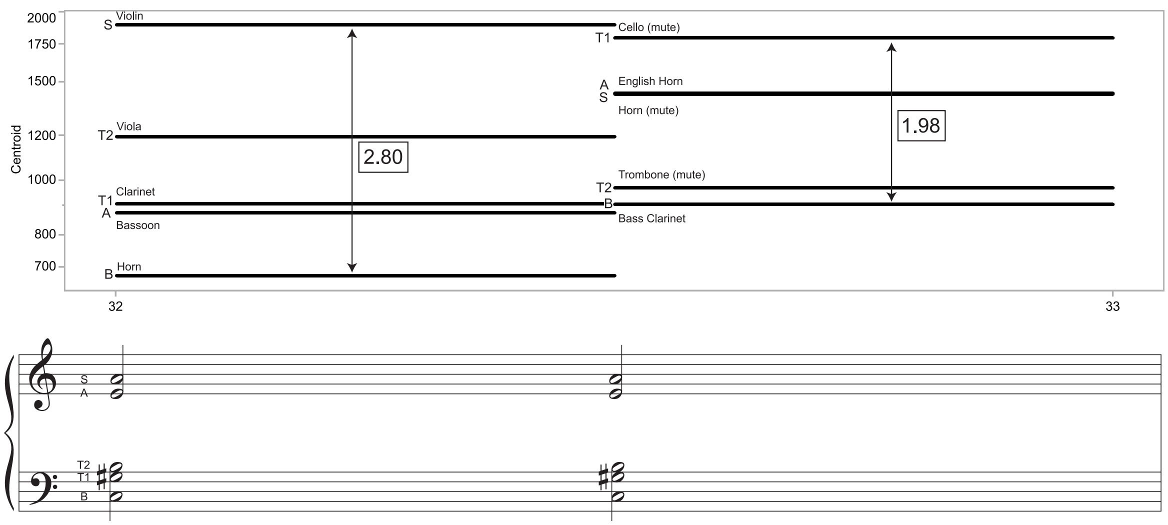


Figure 2-9. “Inverted” span changes, m. 32.

While the foregoing discussion does not attempt a comprehensive analysis of the changing color of the organism in “Farben,” it demonstrates the kinds of analytical claims element-level brightness analysis can give access to. By analyzing the distribution of centroids in these sonorities, we can account for the perceptual effect of instrument changes and their interactions with pitch changes in giving shape to the work in time. In the next section, I

[14] From the instrument-oriented perspective, too, it is worth noting here that the violin takes up the soprano voice in the first chord in m. 32, which can perhaps be seen as a kind of “inversion” of the opening sonority of the piece, where viola and double bass took the bass voice.

develop a complementary approach that highlights the changing colors of each individual voice in the organism.

Sequential Element Contours and Permutations

Thus far, I have limited my discussion to "chord color," primarily focusing on changes in the statistical distribution of centroids within verticalities. But owing to its canonic structure, the piece invites an examination of the individual polyphonic voices (soprano, alto, tenor 1, tenor 2, and bass) and their separate trajectories through the evolving colors of the organism. Zooming in on the changing orchestration of individual voices of the chord unveils a set of brightness contours that go unseen when focusing on verticalities. These contours can be systematically compared to one another, revealing logical relationships between them.

Figure 2-10 deconstructs mm. 13–19 to provide a clear view of the centroid trajectory of the upper four polyphonic voices. In mm. 13–14 (labeled a), the middle two voices are an inversion of one another, moving contrarily in brightness space. Simultaneously, the outer two voices (soprano and lower tenor) do the same, albeit imperfectly. Understanding the outer voices as an inversion of each other may also explain the fact that the muted trombone is doubled by the violin in the third note of the lower tenor line: were it not the case, the resulting centroid would be either too bright (solo violin) or too dark (muted trombone) to mirror the soprano line.[15]

---

[15] It should be pointed out here that the last three notes of the soprano and tenor 2 lines in Figure 2-10a are also "inversions" of each other in terms of instrumentation: the soprano

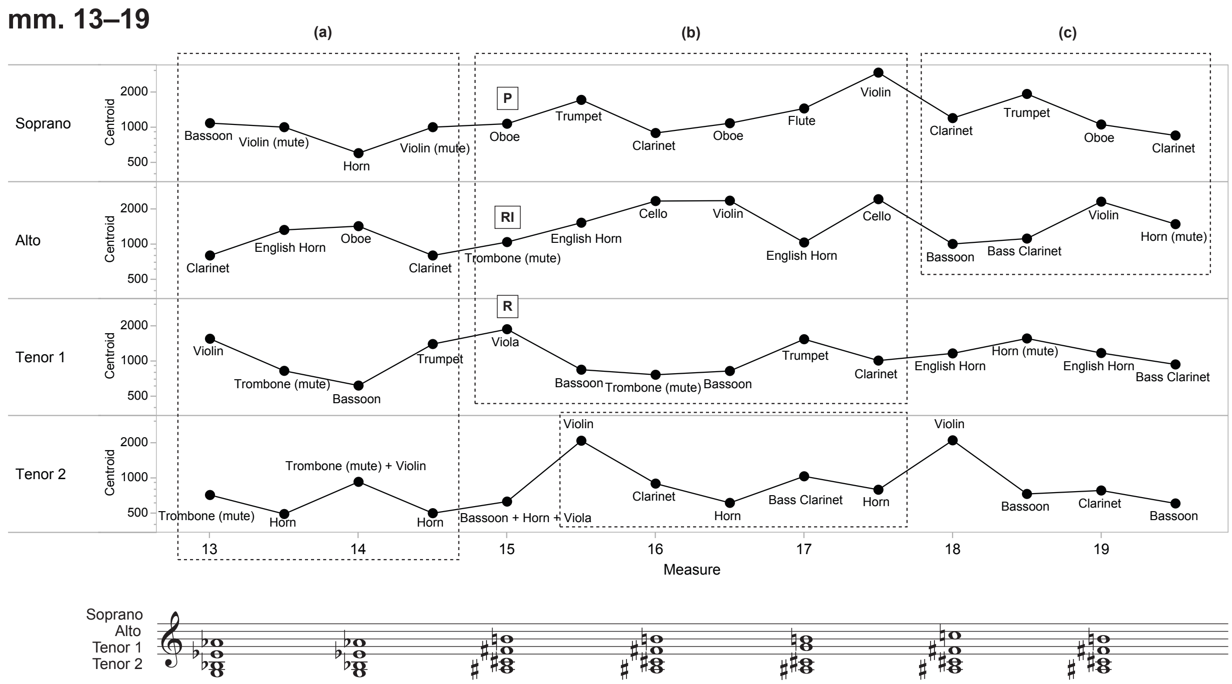


Figure 2-10. Sequential element graph of mm. 13–19 of "Farben"; (a) in mm. 13–14, the alto and tenor 1 are related by inversion, as are the outer voices (soprano and tenor 2); (b) in mm. 15–17, the soprano's brightness contour (P) is doubled by the tenor 1 in retrograde (R), and the alto voice in retrograde inversion (RI); (c) in mm. 18–19 the soprano and alto are related by retrograde.

In the next three measures (b), when T_4 is reached and sustained, the soprano's six-note brightness contour is doubled by the tenor 1 in retrograde, and by the alto voice in retrograde-inversion. The middle voices are thus inversions of each other, continuing the pattern established in the previous measures.[16] In this same section, if we exclude the first

---

line goes "violin → horn → violin," while the tenor 2 line goes "horn → violin (and muted trombone) → horn." This reading, clearly, makes it difficult to account for the muted trombone.

[16] My assignment of "prime" (P) to the soprano line in Figure 2-9b is arbitrary (any of the three voices could be considered the "prime" series).

note of the tenor 2 line, too, its contour from mm. 15–17 resembles the retrograde of the soprano series (i.e., tenor 1), albeit compressed to span five notes instead of six. In the next section, measures 18–19 (c), the alto line is in retrograde of the soprano line. It is notable that in all these examples, looking at instrumentation alone would not reveal any well-defined pattern: the soprano line in Figure 2-10c, for instance, progresses through clarinet, trumpet, oboe, and back to clarinet, while the alto line involves none of these instruments (and lacks a repeated instrument that would correspond to the clarinet). Focusing on the brightness of the lines, nevertheless, clearly reveals a pattern in the alto line that reverses that of the soprano.

During the final section of the work, as the canon returns in inversion, several element contours from the middle section return in permuted form. In mm. 33–36, the soprano voice features a series of seven tones that are an exact retrograde of an earlier seven-tone passage in the same voice (mm. 15–18). Figure 2-11 displays these two sequences, with numbers to show the corresponding tones in each.

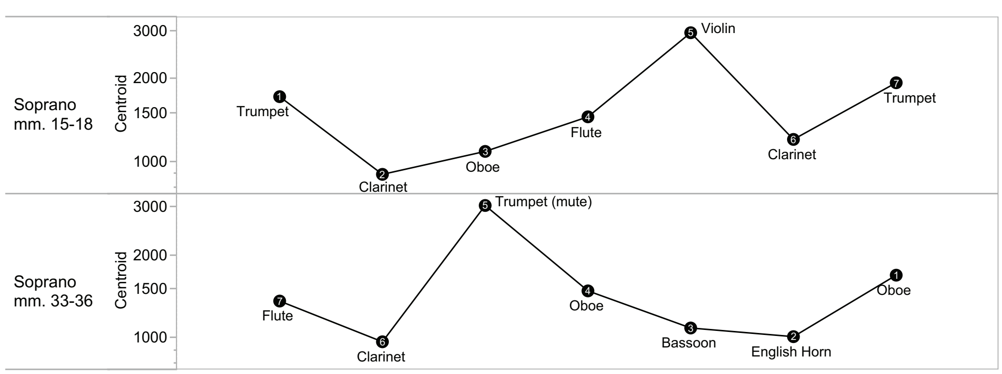


Figure 2-11. A seven-tone sequence of element-level centroids (top panel) returns in the final section of the work in retrograde (bottom panel).

During the same passage (mm. 33–38), large sections of both tenor voices' element contours are permutations of earlier tenor sequences found in the middle section. Beginning in measure 33, the tenor 1 voice, remarkably, features a twelve-note series that is a near-perfect retrograde-inversion of a sequence from earlier in the tenor 2 voice (mm. 13–18). Figure 2-12a shows both contours, with numbers matching each tone to its permuted equivalent. The tenor 2 voice in mm. 32–38, too, is a retrograde-inversion of the tenor 1 voice from the same earlier passage (mm. 13–19), albeit with slightly more flexibility. Figure 2-12b displays the two sequences in question, with colors to clarify the particular segments of the initial series that observe the retrograde-inversion. Notably, a four-note stretch in the middle relates to the original series by inversion, deviating temporarily from the retrograde-inversion.

That these permutations should occur in the final section of the work, in tandem with a dramatic turning point in the pitch domain (i.e., the inversion of the canon), points to a deep and yet unexplored interactivity between brightness contours and pitch structures in "Farben." In the aftermath of the dizzying climax from mm. 24–30, it is as if both pitch and timbre are incited to alter their approach and revise earlier passages: the canon is inverted, and timbral paths are retraced and refined. Surely, many more patterns like these can be detected throughout the piece, which will impact our understanding of the nature of this interactivity.[17] But rather than attempt an exhaustive analysis here, my intention is instead

[17] Analysis using multiple sample libraries, too, will surely reveal further patterns and, plausibly, extend the boundaries of those I've already found.

to open a line of inquiry based on the understanding that meaningful patterns are found in sequences of element-level centroids, and that these structures would go undetected if we concerned ourselves solely with the work's instrumentation.

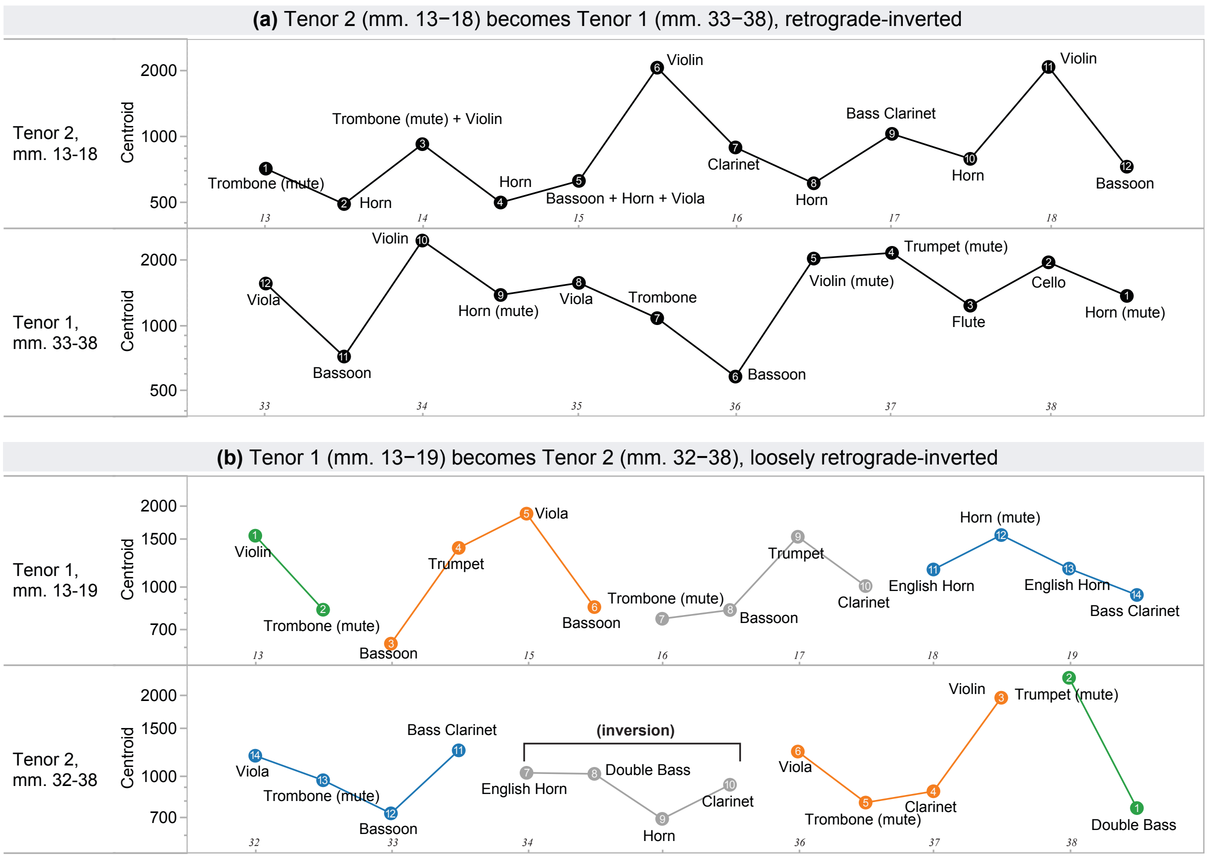


Figure 2-12. (a) The entire tenor 2 brightness contour from mm. 13–18 recurs in retrograde inversion in the tenor 1 voice in mm. 33–38. Numbers match each tone to its RI equivalent in the permuted sequence. (b) Segments of the tenor 1 part from mm. 13–19 appear in retrograde inversion in the tenor 2 part in mm. 32–38. Colors match the corresponding segments to each other. The grey segments in the middle are related by inversion rather than retrograde inversion.

In the final section of this chapter, I want to counterbalance this zoomed-in approach by expanding my focus to the level of the layer and sketching a path toward a textural understanding of the piece. This will involve examining a handful of sonic interactions between the organism, with which we've been solely concerned thus far, and the often-neglected foreground elements of the piece.

Sketch of a Layer-Level Analysis

As I mentioned previously, perhaps the most perceptually meaningful way to discuss brightness in the "Farben" chords is to take a single centroid value for each sonority that reflects the overall brightness of the organism at any given moment. Importantly, this overall brightness value is not equal to the average element-level centroid of the sonority. This is because the overall layer-level centroid takes into account the relative loudness of each tone in the mix: while all tones may be marked with the same dynamic level, their natural carrying powers result in some instruments masking others or otherwise gaining a larger share of the overall signal.

Figure 2-13 below shows a layer-level centroid analysis of both the background and foreground strata of mm. 15–25 of "Farben." Centroid data in the top panel coordinates with pitch material in the bottom panel via color. Thus, connected black points along the bottom of the upper panel represent the changing overall brightness of the organism. From this vantage point we can detect an overall darkening trend in the organism—as its pitches gradually inch downward over the course of the short passage, its overall timbre likewise becomes ever darker (before suddenly brightening up again at measure 25 with another string punctuation). In spite of this trend, we can clearly see brightness peaks in the organism throughout, and beginning in the second half of measure 17 these peaks occur at regularly spaced intervals of three half notes.

The passage appears to divide into two halves (labeled a and b in Figure 2-13), each of which is nearly symmetrical in some way. In (a), the organism brightens in the second half of measure 15, then darkens again, as if to make room for the piccolo and celesta figure (red), which doubles the top note of the organism in mm. 16–17. When this added layer ceases, the organism brightens itself again as if to make up for the lost brightness in the second half of measure 17. In measure 19 the organism reaches a peak again and darkens considerably to be as bright as it was when it began in measure 15—yielding a long sequence of centroids that appears nearly symmetrical (indicated with dashed ovals in Figure 2-13a).

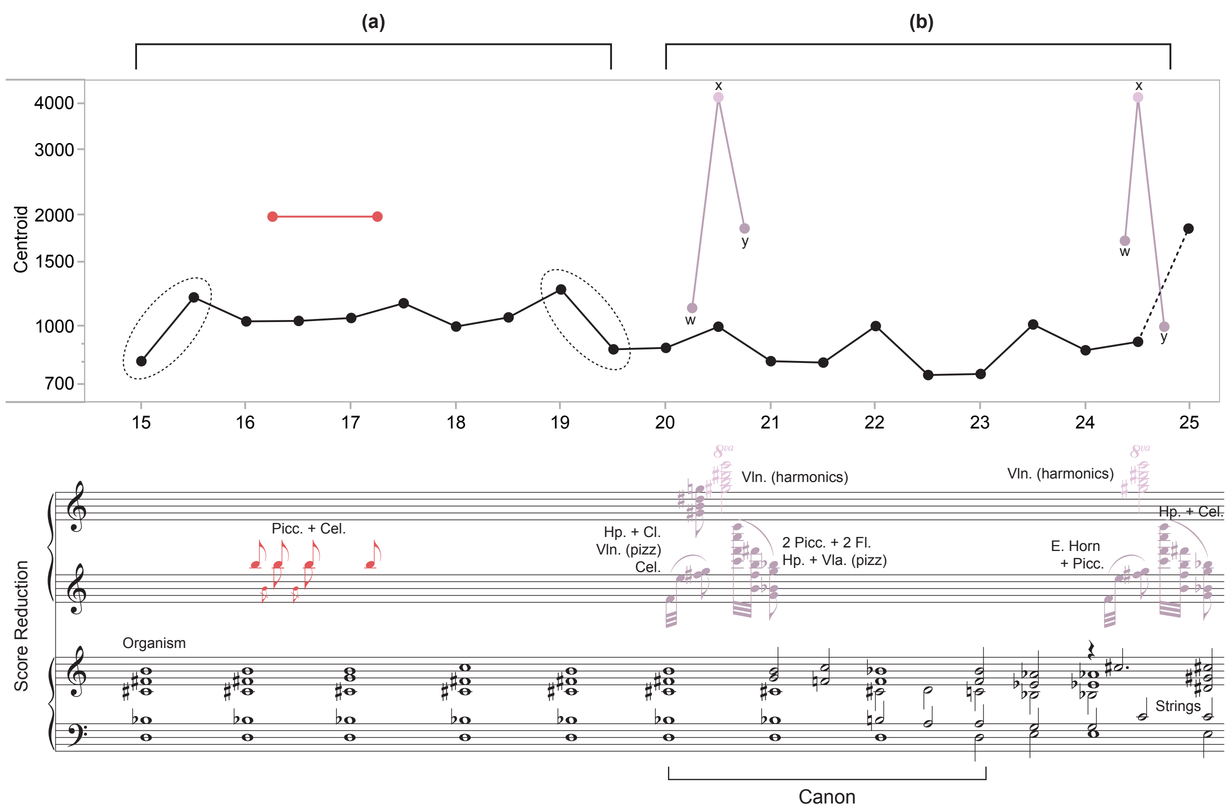


Figure 2-13. Concurrent layer graph of mm. 15-25. The organism's overall centroid inches downward in tandem with pitch. (a) and (b) are both nearly symmetrical timbral structures.

In (b), a distinctive foreground gesture (highlighted in purple) appears twice, in m. 20 and 24, with different orchestrations. Both instances of the foreground gesture divide into three motives, each of which is given its own centroid in Figure 2-13b. Following Burkhart, I label these motives *w*, *x*, and *y*. In both cases of the gesture, the *x* motive is played by extremely bright violin harmonics. The *w* motive (the so-called "leaping trout") appears first as a blend of harp, clarinet, and celesta in m. 20, before returning in m. 24 as a *brighter* combination of English horn and piccolo. Simultaneously, the piccolo and flute statement of *y* in m. 20 returns in a *darker* color (celesta + harp) in m. 24. Thus, *w* and *y* appear to switch places in terms of brightness, yielding a kind of timbral symmetry that bookends a statement of the canon.

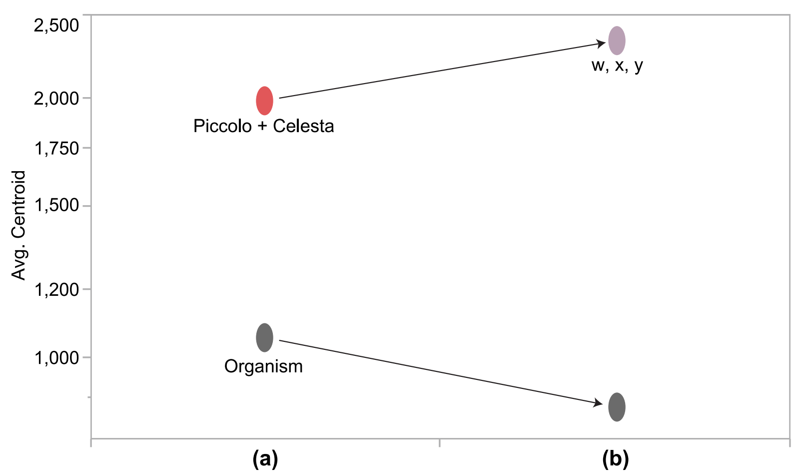


Figure 2-14. In mm. 15–25, the foreground material becomes brighter as the organism becomes darker.

A major benefit of the layer-level approach is its ability to estimate the degree to which layers do or do not overlap in spectral space. In Figure 2-14a, the piccolo and celesta layer is wholly distinct from the organism, as evidenced by the large ratio that obtains between them. And, over the course of the passage, the added foreground material becomes on the whole brighter, while the organism becomes darker, increasing the concurrent centroid ratio between layers and the resulting distinctness of each layer (Figure 2-14).

This brief sketch outlines the kind of analytical observations that can be made by examining centroids and their interactions at the level of the layer. A more comprehensive analysis of these interactions through the piece falls beyond the scope of this chapter but would surely reveal timbral patterns that influence our understanding of the structure of the work. The next chapter, on Ruth Crawford's *Music for Small Orchestra,* builds further on this outline in a large-scale analysis of texture via layer-level brightness.

While a complete analysis of the timbral structure of "Farben" cannot be fully pursued here, it seems to me that a multilevel reading that progresses through all three textural levels—*element*, *layer*, and *texture*—would align well with the piece's unfolding drama. In the opening section, regular changes in orchestration combine with a clear canonic procedure to draw us in, inviting us to scrutinize the timbres of individual elements. As additional layers are introduced, the focus shifts, and we view the organism as a unified entity distinct from the splashes of color in the foreground. At the climax of the piece, when both the pitch canon and orchestration accelerate and the foreground layers intensify, we respond to the increasing complexity by zooming out even further, attending to the overall brightness of the entire texture. On this reading, the last section of the piece then reverses the process, encouraging a renewed focus on elements and their subtle transformations as the piece draws to a close.

Conclusion

This chapter focuses on the potential of analyzing timbre via concurrent and sequential element-level centroids in “Farben.” From a discussion of the color composition of opening sonorities emerges a bundle of salient element-level features that can be traced through time to reveal meaningful structural patterns. The individual voices of the “Farben” organism produce sequences of centroids that interact with and permute each other in apparently logical and powerfully suggestive ways. The entire organism as a layer changes its brightness in response to the behaviors of other layers in the texture. Overall, the centroid-based approach developed in this chapter allows us to see and hear things that we wouldn’t otherwise see or hear—subtleties that go unnoticed when focusing on instrumentation or spectrographic approaches to timbre.

Chapter 3

Visualizing the Relative Brightness of Concurrent Textural Layers in Ruth Crawford's *Music for Small Orchestra* (1926)

Quantifying the relative brightness of simultaneous instrumental sounds can give us a better understanding of timbre's role in the perceptual organization of texturally complex post-tonal music. Scholars of musical stratification, however, have converged on a definition of texture that privileges rhythmic coordination. Proponents of "partitional" approaches to texture (Gentil-Nunes, 2017; 2020; Guigue & de Paiva, 2018; Moreira, 2019) understand polyphonic music to be divisible into layers (or "parts") primarily on the basis of onset synchrony and shared contour as represented in the symbolic data of the score. This chapter argues that this model is unsatisfactory for characterizing musical texture unless supplemented by empirical assessments of brightness. Through an analysis of Ruth Crawford's *Music for Small Orchestra* (1926), I show how brightness measurements allow us to assess relationships both within and between rhythmically coordinated layers as defined by partitional analysis. To facilitate my analysis, I develop a new kind of partitional analysis based on timbre which I call a *transparency graph* and address some of its analytical implications.

Partitional Approaches to Texture

Theoretical studies focusing on texture have proliferated in recent years. Within this literature, a prominent trend is the "partitional approach," which draws significant influence from Wallace Berry's discussion of texture in Chapter 2 of *Structural Functions in*

Music (1976). As part of a larger treatise on the "structural" elements of score-based music, Berry lays out an approach to the analysis of written textures, defining three parameters, or "factors," that govern whether concurrent elements will adhere into a unified layer or remain segregated as independent layers: the *rhythmic, directional,* and *intervallic* factors.[1] In Berry's approach, and in the partitional theory based on it, the most important of these three factors in determining integration appears to be rhythmic alignment. And, while Berry allows for what he calls "coloration" (i.e., timbre and orchestration) to impact the analytical classification of textures, it does not constitute a proper "factor" in the system, but rather serves as a complicating element—a consideration that can either weaken, strengthen, or otherwise problematize a largely rhythmic and/or pitch-based textural analysis.[2] A major goal of this chapter is to extend and refine the analytical framework initiated by Berry, elevating in particular the role of timbre.

A particularly useful feature of Berry's approach is his numerical system for symbolizing a textural configuration—i.e., the number of concurrent layers present, along with the number of elements in each layer. For example, music containing six elements (or

---

[1] Each of these three parameters is modified with one of three added prefixes—*homo-hetero-*, or *contra*—to denote sameness, mild variation, or pronounced contrast, respectively, yielding nine texture descriptors.

[2] This approach closely resembles a later theory developed by David Huron (1989), who posits two sufficient factors: *onset synchrony* (i.e., rhythmic coordination) and *semblant motion* (i.e., similar and parallel motion). More recently, Ben Duane (2013, 2017) proposes four factors, which he calls onset and offset synchrony, pitch co-modulation, and harmonic (i.e., overtone) overlap.

"sounding components," in Berry's parlance) might be grouped into unified layers (or "real components") in a number of different ways. Figure 3-1 below shows three such possibilities, using Berry's numerical notation. Figure 3-1a represents six elements grouped into three independent layers, one containing three coordinated elements, another containing two, and a third containing a single element. Figure 3-1b shows the same total number of elements now grouped into two independent layers (of four elements and two elements, respectively). 3-1c presents the six elements coordinated in a single layer. The vertical ordering of integers loosely corresponds to register of the layers.

$$\text{a)}\ \begin{array}{c}3\\ \hline 2\\ \hline 1\end{array} \qquad \text{b)}\ \frac{4}{2} \qquad \text{c)}\ 6$$

Figure 3-1. Three configurations of six elements in Wallace Berry's textural notation.

Berry never uses the mathematical term "partition," but his approach here epitomizes the concept, notating how an integer (such as 6, in this case) can be represented as a sum of smaller integers, or "integer parts."[3] Indeed, in recent years there has been a notable surge of math-oriented work that explores the implications of Berry's partitional approach (Gentil-Nunes, 2017; 2020; Guigue & de Paiva, 2018; Moreira, 2019). This literature employs a more standard partitional notation, written as $x(a, b, c \ldots)$, where x is the total number of elements in the texture, and each of the parenthesized parts (a, b, c) is a layer of an indicated size—a subset of the total that belongs together by virtue of being

[3] For an in-depth study of the mathematical theory of integer partitions, see Andrews (1984).

rhythmically coordinated in the score. Figure 3-1a above is thus rewritten as 6(3,2,1), indicating a six-element texture partitioned into three layers (of varying size). The sum of the parts $(a + b + c \ldots)$ is always equal to the whole (x). The greater the number of parts (i.e., the cardinality of the partition), the more simultaneous, rhythmically independent layers are present in the texture. For example, a long string of six ones—6(1,1,1,1,1,1)—symbolizes a texture made up of six rhythmically independent strands, each containing one element. These long strings are written with a conventional shorthand using a superscript numeral: $6(1^6)$.

A detailed review of this partitional literature is beyond the scope of this chapter, but it will be sufficient to demonstrate that these authors, despite admitting a modest role for perception, ultimately leave timbral interactions out of the picture. Although this score-centric, coordination-based textural analysis offers a deeply informative starting point, I argue that it remains unsatisfactory for modeling the character of a musical texture unless supported by an empirical assessment of timbre obtained through acoustic measurement. Before delving into my approach to this kind of timbre-oriented analysis of texture, it will be beneficial first to consider some useful facets of the partitional strategy in conversation with a piece of texturally (and timbrally) complex post-tonal music. The central case study here is Ruth Crawford's *Music for Small Orchestra* (1926), a piece whose rhythmic organization suggests a straightforward textural partitioning, but whose timbral dimension problematizes such an approach.

Textural Partitions in *Music for Small Orchestra* (1926)

Figure 3-2 shows a condensed score of mm. 3–9 of the piece with a partitional overlay. The piece begins with the pitch F4 repeated nine times in the piano (not shown). When the wind and string instruments join in the third measure the music segregates into two distinct layers—eight instruments sustain a dissonant six-note chord, while the piano continues its established pattern. In the next measure, two violins break off from the sustained chord, and in unison alternate between F♯4 and E4 (the two pitches on either side of the piano's F). Thus, after having established a two-layer configuration, the texture undergoes a subtle transformation: the nine components, initially grouped into eight plus one, become grouped into six plus two plus one.

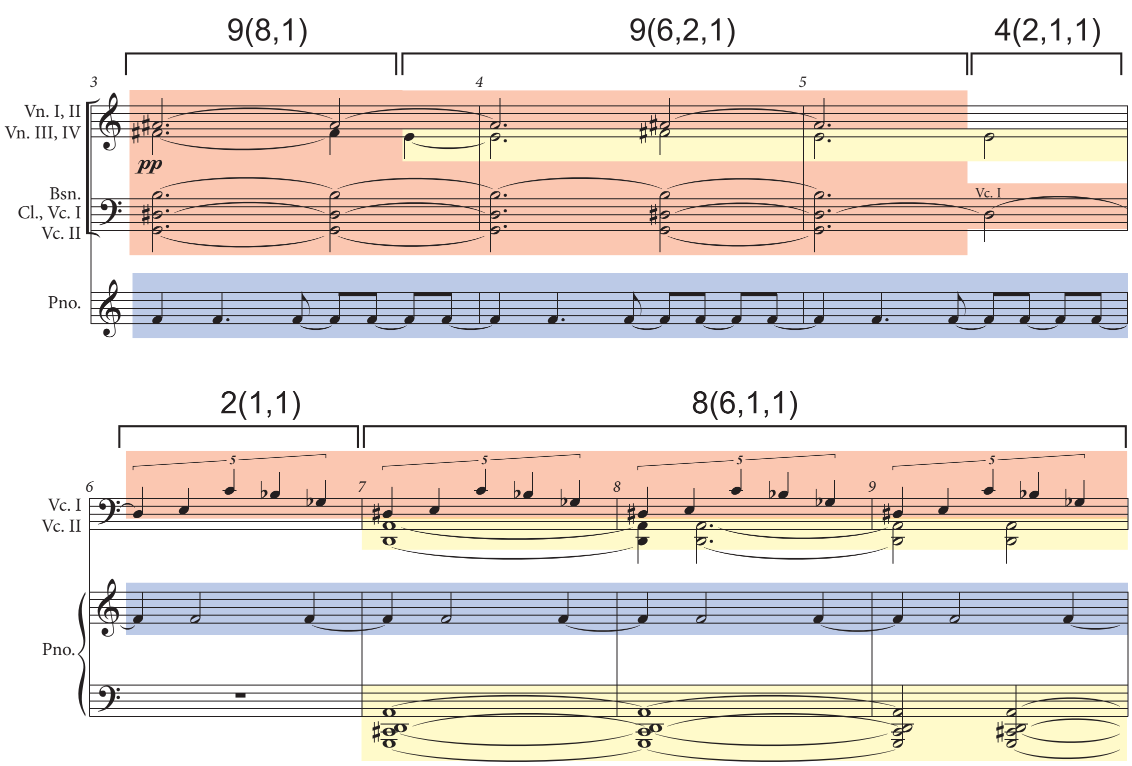


Figure 3-2. Partitional score annotation of mm. 3–9 of Crawford's Music for Small Orchestra. Nine elements group into two layers (8,1), and then three layers (6,2,1). In measure seven, a very different texture receives a similar partitional analysis (6,1,1).

The two violins are playing in unison, but I've labeled them here as two separated entities (e.g., with the number 2). In fact, it is not exactly clear how unison and octave doublings should be accounted for in partitional analysis, and no consistency is found in the literature. A unison-doubled line can either be viewed as (a) multiple elements maximally coordinated, or (b) a single element with a composite timbre.[4] On the one hand, it seems logical to include all sounding instruments in a textural analysis, but on the other it seems unsatisfactory that, for example, three instruments in perfect unison and three instruments combining to form a triad should yield the same partitional expression—3(3). In the present study, I follow (a) above, viewing unisons as maximally coordinated multiples, but to add further clarity where necessary, I will employ nested partitions in square brackets to show the pitch-based sub-partitioning within each coordinated part. The cardinality of these sub-partitions is equal to the number of unique pitches within the layer, and the value of each of the sub-parts indicates the number of instruments producing a single pitch. For the example of a trio playing in unison vs. a triadic configuration, this nested notation will show the same *overall* partition, 3(3), but different pitch-based *sub-partitions*—3(3[3]) for unison, and 3(3[1,1,1]) for triads—thereby encoding both a salient similarity and a crucial difference. The interior partitioning of the layers shown in Figure 3-2 is thus clarified in Figure 3-3 below.

[4] Moreira (2020) makes the distinction between these two readings based on score layout, i.e., choose (a) above if the unison is written on a single staff (and marked, say, *a2*), and choose (b) if spread across multiple staves.

$9(8[2^3,1^2],1)$ $9(6[2^2,1^2],2[2],1)$ $4(2[2],1,1)$ $2(1,1)$ $8(6[2^2,1^2],1,1)$

Figure 3-3. Partitional reading of mm. 3–9, revised to account for unison doublings. Subpartitions in square brackets show how layers are grouped by pitch.

In measure six, the first cello introduces its crawling five-note ostinato, which is then offset by a low, gong-like four-note chord in the piano, of which two notes (the fifth D2–A2) are doubled by the second cello. This texture sounds and looks decidedly different from the opening one, but the partitional reading yields a very similar sequence of numbers: eight elements grouped into six plus one plus one. While it fails to capture the sense of change, the similarity of the partitional analyses does encourage us to read associations that we wouldn't otherwise: associations between, for example, the group of six sustained notes in the strings and winds in the opening texture, and the six gong-notes in the second one. These are two layers of equal size (six elements), each serving a similar role in a texture containing three layers. And, as the unison-inclusive partitions show in Figure 3-3, both chords contain six elements, but only four distinct pitches: in the first chord, the perfect fifth D#3–A#4 is doubled, while in the second chord, the perfect fifth D2–A2 is doubled.

Quantifying Coordination: Agglomeration and Dispersion

Working with broad indications of the number and size of layers allows for useful summaries to be made about how texture unfolds in time. In the partitional literature, a favorite summarization tool, developed by Gentil-Nunes (2009), involves assessing the degree to which a texture is coordinated into layers.[5] For a given partition, we can calculate two important values: *agglomeration* (or "interdependence," for Berry), which indicates how coordinated the overall texture is, and *dispersion* ("independence"), which measures layer non-coordination.[6] Together, these values tell us about the relative independence of parts: a chorale will have high agglomeration values, while a fugue will have high dispersion values. To get these values we need to know how many *pairs* of elements are rhythmically coordinated, and how many are not. Thus, we begin with the "n choose 2" formula, which calculates the number of pairwise combinations (C) from some number of elements (n):

$$C_n = \frac{n(n-1)}{2}$$

The *agglomeration* index, then, is the sum of all pairwise combinations (C) calculated for each discrete layer in a given partition or texture. Given the partitional expression $x(a, b, c)$ the agglomeration index is equal to $C_a + C_b + C_c$. A high agglomeration index thus indicates that many pairs of elements are grouped together into rhythmically unified layers. *Dispersion* is agglomeration's opposite; it tells us how many pairs of elements are *not* coordinated with each other. First, we find the total number of possible pairs in the texture, and subtract the agglomeration index, leaving the number of *unconnected* pairs.

---

[5] Pauxy Gentil-Nunes (2009, 2017, 2018) is a central figure in the mathematical development of Berry's theory. His work, which he calls "a radical expansion" of Berry's approach, fleshes out many of the mathematical implications of Berry's framework.

[6] "Agglomeration" and "dispersion" are terms used in statistics, but Gentil-Nunes' usage here is wholly distinct.

For the partitional expression $x(a, b, c)$, the dispersion index is equal to $C_x - (C_a + C_b + C_c)$. These two indices—agglomeration and dispersion—can be charted through time and visualized on a combined *coordination graph*, thus capturing a salient quality of textural organization.[7]

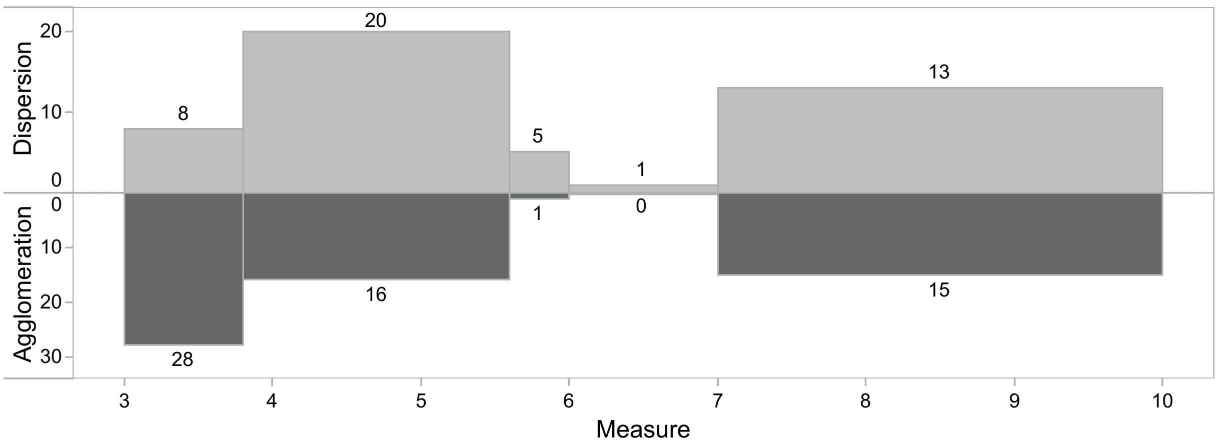


Figure 3-4. Coordination graph of mm. 3–9 of Music for Small Orchestra, showing an alternation between agglomerative (bottom) and dispersive (top) tendencies.

To illustrate, let us return to initial partitional analysis of *Music for Small Orchestra* given previously in Figure 3-2. This analysis is summarized in the mirrored bar-graph in Figure 3-4, with dispersion indices shown along the top panel and agglomeration indices along the bottom. This kind of visualization represents a benefit of the partitional approach, in that it allows us to readily perceive changes in the overall nature of the texture: in this instance, it reveals a dynamic interplay over time where the musical material alternates between

---

[7] Gentil-Nunes refers to this kind of graph as an "indexogram," given that it plots the two indexes over time. Guigue and de Pavia refer to "relative voicing complexity" graphs. I prefer to use the simple term "coordination graph" here.

being (a) consolidated into thick individual layers (indicated by high agglomeration indices) and (b) spread across multiple distinct layers (indicated by high dispersion indices).

In this push-and-pull between agglomeration and dispersion, dispersion ultimately prevails. After the cello's quintuplet ostinato is established (mm. 7–10), a solo flute and bassoon are added to the mixture, each playing an independent line. The process continues with the addition of the clarinet (m. 14), followed by two pairs of violins (mm. 19 and 25), all rhythmically isolated and unblended. In measure 25, the piano moves away from its repeated note to a variation of the cello quintuplet in parallel octaves, and by measure 33, the music has transformed into a dense and dispersed tapestry of sound, with sixteen elements grouped into eight layers of different sizes. This first section of the work (mm. 3–44) thus enacts an additive process of snowballing dispersion, continually introducing to the texture new isolated elements that each remain rhythmically differentiated and unintegrated.[8] Figure 3-5 shows a coordination graph for the entire first movement of the work, highlighting this accumulative process.

[8] In this piece, Crawford's preferred technique for rhythmically distinguishing multiple concurrent elements is clearly to use multilayered polyrhythms, which entails various distinct subdivisions of a common whole note, in a manner described by Henry Cowell in Part II of *New Musical Resources* (1930/1997, pp. 45–66).

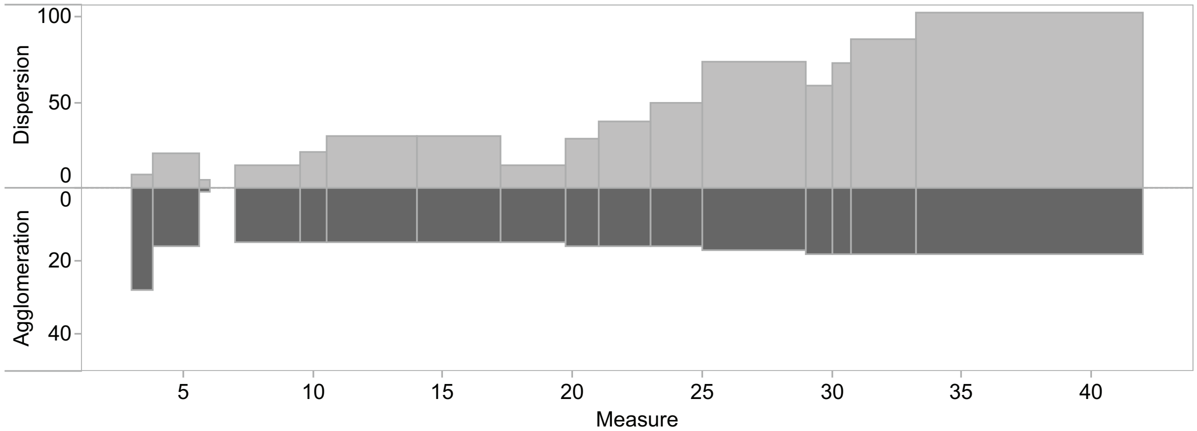


Figure 3-5. Accumulating dispersion in mm. 1–44 of *Music for Small Orchestra*.

Music for Small Orchestra, with its strict rhythmic delineation of layers, seems an almost perfect case study for score-based partitional analysis. As I've suggested already though, there seem to be important sonic relationships that obtain between these layers, and within them, that are not accounted for by an approach based on symbolic score information. Let us discuss two pertinent examples.

Figure 3-6a shows the duo of flute and high-register bassoon that joins the texture in measure 10. While these instruments remain more-or-less rhythmically isolated from one another throughout the first section of the work—a few passing alignments notwithstanding—there is a definite acoustic similarity between them. This timbral affinity, often pointed out in orchestration manuals and exploited by composers, bears on our perception of layer interaction, but goes unseen in a coordination-oriented partitional analysis.[9] For a textural analysis to represent perception more adequately, it must have some method by which interlayer (dis)similarities can be modeled. Layer-level brightness analysis points to one such method.

Figure 3-6. (a) interlayer similarity in mm. 10–14 of *Music for Small Orchestra* (flute and bassoon only); (b) intralayer dissimilarity in mm. 56–59 (woodwinds and piano only).

Timbral interactions *within* layers can be modeled too. Figure 3-6b isolates a moment from later in the piece (mm. 56–59), when three wind instruments and piano take up a sextuplet ostinato in close five-voice homophony. One might justifiably challenge a partitional reading that says these elements cohere into a single layer—$5(5[1^5])$. While it certainly does so on the page, can we hear it as an integrated unit? Is it more likely to resist integration and

[9] For perceptual similarity of flute and bassoon, see Brant (2009, p. 56), who includes the high-register bassoon (below a mezzo-piano dynamic) as part of his Wind Group I (a group of “flute-like” timbres).

remain as five separate elements, or perhaps adhere to some other configuration, like the two piano notes against the wind trio? There appears to be a high degree of intralayer dissimilarity here concealed by the partitional approach, which might be teased out by an analysis of element-level brightness within the layer. In these two examples, rhythmic partitioning is in tension with auditory grouping: uncoordinated layers seem to cohere (Figure 3-6a), and coordinated layers seem to remain segregated (Figure 3-6b). The partitional approach conveys something important about the organization of musical material, and thus serves as a useful starting point for texture analysis. However, to effectively model the kind of polyphonic interactions just described, it becomes necessary to move beyond the notated score and into to the domain of acoustics.

In the section that follows, I present detailed acoustic analyses of interlayer relationships within the piece using layer-level centroids. I then explore briefly how intralayer relationships can be modeled using concurrent element-level centroids, which we covered in Chapter 2. Through a reappropriation of partitional thinking, I develop the dual concept of *clustering* and *transparency,* which I use to summarize the interactions of layers through time.

Interlayer Brightness

As described in Chapter 1, analyzing the relative brightness of concurrent textural layers involves defining concurrent layers and textural units through score-based partitional analysis, as described above. Then, for each textural unit, I use MIDI data realized by a sample library to isolate each layer and export the audio playback. Average centroid values are calculated for all sound files in the textural unit and plotted on what I call a *concurrent layer graph*, with selected SCQs labeled. Each of the points in a concurrent layer graph thus represents a rhythmically delineated textural layer, as defined by partitional analysis. Each point's height along the *y* axis corresponds to its brightness.

To illustrate, let us revisit the three-part texture that begins in measure 3 (shown above in Figure 3-2). As described previously, the texture consists of the solo piano's repeated F4, two violins in unison, and the six-part chord of winds and strings. With the score of this excerpt encoded in MIDI data and loaded into a MIDI sequencer (MuseScore), I am able to isolate the three layers from one another and listen to each one performed separately by a sample library (MuseSounds). I then export each of these synthetic performances as a separate sound file and calculate the average SC value of each file, along with the SCQ values between the files. Figure 3-7 below displays (a) the average SC of each layer in the opening texture, and (b) the SCQ values between them. In the bottom panel (c), this data is visualized in a concurrent layer graph, with the largest SCQ labeled (4.23).

(a) Centroids (Hz)

Layer	
solo piano	555
two violins	2,346
six-part chord	1,067

(b) SCQs

Layer Pairs	
solo piano vs. six-part chord	1.92
solo piano vs. two violins	4.23
two violins vs. six-part chord	2.2

(c) Concurrent Layer Graph

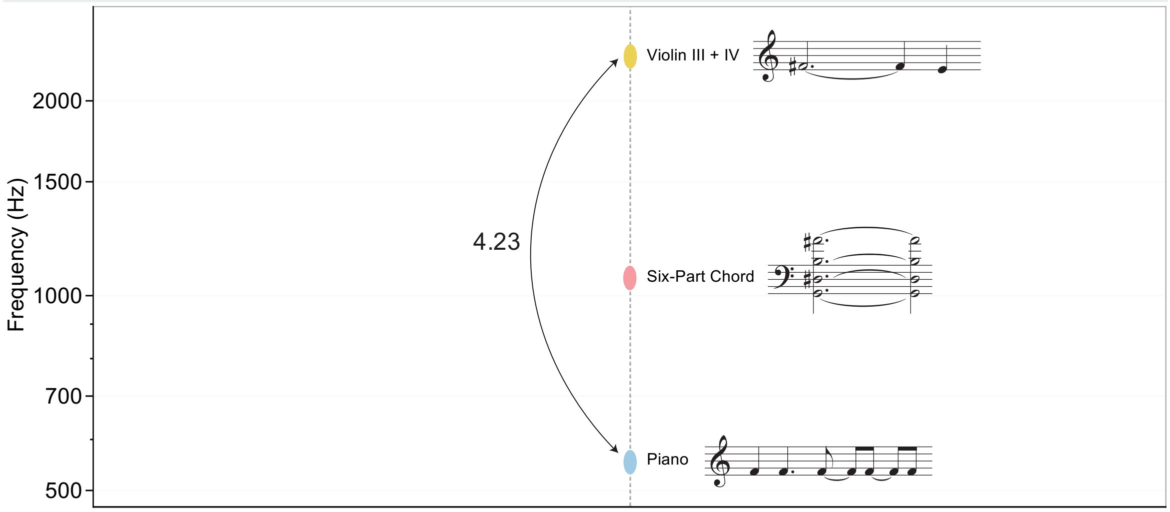


Figure 3-7. (a) Centroid data; (b) SCQ data; (c) Resulting brightness analysis of concurrent layers in the first textural unit (mm. 3–5). The six-part chord falls between two extremes spanning a high SCQ (4.23).

Reconsidering this opening texture now in terms of the relative brightness of its three layers, we find that the violins and the piano, although they occupy the very same region in pitch space, are completely separate in brightness space (the violins are roughly four times brighter than the piano). The six-part chord, despite containing both higher and lower pitches, falls almost precisely in the middle of these two layers, bridging the timbral gap between them. In terms of pitch space, the solo piano sits near the top of the sonority, but nevertheless appears to be the darkest layer in the texture. What was suggested in Chapter 2 remains applicable here: I proposed that centroid distributions, like the one illustrated in Figure 3-7, can effectively capture a form of "perceptual spacing" among elements in a chord, or in this case, layers in a texture. I submit that when listening attentively to this opening passage, it is not overly challenging to perceive two violins sitting "at the top" of the texture, with the solo piano positioned "at the bottom." While conventional music theory takes as given that distribution in pitch space represents a perceptual ordering of simultaneous sounds, it seems clear that another very real set of perceptual relationships is captured by the centroid distribution, which very often does not align with pitch.

Figure 3-8 shows a concurrent layer graph that includes the two textures immediately following this passage (Units 2 and 3). Like the first one, the second major textural unit is also divided into three layers: the cello ostinato, a low-pitched "gong" chord, and a repeated F4 in the piano. Surprisingly, the gong chord, played on the piano and doubled by the cello, is not the darkest sound in most of these textures. Owing to the bright, almost metallic attack of the low piano notes, and the prominence of the noisy bow sound in the cello, it consistently outshines the mellow repeated F in the right hand of the piano (shown in light blue), which is typically the darkest layer in each texture. The arrangement of the layers in the second texture is thus very similar to that of the first but is now compressed into a smaller overall space: two outer layers (piano and cello) are related by an SCQ of 1.83, and the timbral gap between them is bridged by an intermediary layer (the gong chord). As we saw previously, the partitional analyses of these first two textural units invites us to read a similarity between the six-part chord in Textural Unit 1 and the gong chord in Textural Unit 2. Indeed, the centroid distribution of the two textures further supports such a reading, as both chords function analogously as spectral intermediaries between two more distantly related layers.

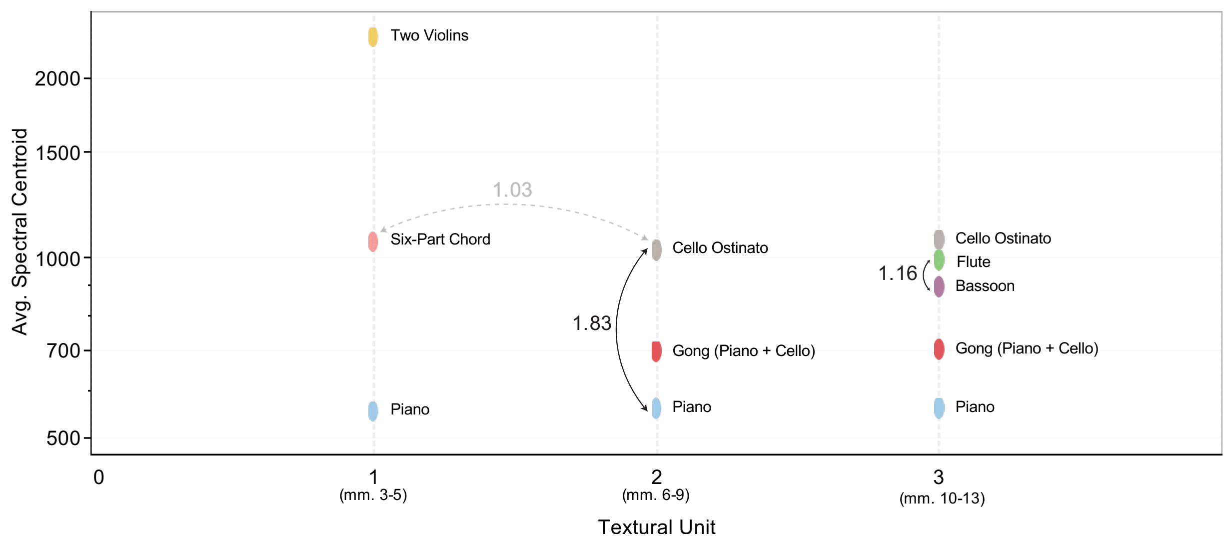


Figure 3-8. Concurrent layer graphs of the first three textural units of *Music for Small Orchestra*. Textural Unit 2 spans a smaller SCQ than Textural Unit 1. The low-pitched gong chord in Textural Unit 2 functions as a spectral "intermediary," much like the six-part chord in Textural Unit 1. The flute and bassoon in Textural Unit 3 form a cluster with the cello ostinato.

The centroid analysis in Figure 3-8 reveals another timbral connection partially concealed by the partitional analysis: the six-part chord in the first unit and the cello ostinato in the second unit are related by an extremely low sequential SCQ (1.03), indicating a subtle perceptual continuity between the two layers (labeled with a dashed grey arrow). Indeed, it is possible to hear the six-part chord as carving out a distinct auditory space that the cello ostinato later fits into. This association is supported by the fact that the cello's ostinato emerges from a held D♯3, which begins its life as a member of the six-part chord in Textural Unit 1.[10]

---

[10] It should be noted here that the placement of both the six-part chord and the cello ostinato on the concurrent layer graph reveals a limitation of the averaging technique described in Chapter 1. Both layers span a wide range of pitches, producing sounds both brighter and darker than the mean value presented in the graph. An element-level analysis

In Textural Unit 3, the flute and bassoon are introduced, but little else about the texture is changed. The previously noted timbral affinity between this pair is represented in the concurrent layer graph as a low SCQ (1.16): they occupy the same region in brightness space, reflecting the fact that they are more likely to overlap perceptually. The graph also reveals that their average centroids are both in close proximity to that of cello ostinato. Indeed, I view the flute and bassoon layers as *emerging from* the cello figure, as a kind of timbral outgrowth that remains rhythmically independent but sonically aligned enough with the cello to form an interconnected cluster of three layers. Thus, while a partitional reading might have it that the third textural unit contains five equally distinct layers (i.e., two more than Textural Unit 2), the brightness reading clarifies that these five layers are arranged into three groups or "clusters" of layers. This perspective thus highlights an important connection between Textural Units 2 and 3: although they feature a different number of layers, they share the same number of layer clusters (i.e., 3).

Clustering and Transparency

Let's delve deeper into the notion that layer centroids can form clusters. Perceptually, a cluster of layers is a group of independent lines that overlap and/or occlude one another, whereas non-clustered layers remain perceptually segregated.[11] This has implications for

---

would thus be necessary to reveal the timbral spread of each layer's constituents. This kind of analysis will be undertaken later in this chapter.

[11] It is possible, also, that layer overlap may result not in occlusion but in perceptual *fusion*, wherein rhythmically independent lines are no longer heard as distinct auditory streams. On my reading, this does not occur in *Music for Small Orchestra*.

the analysis of orchestration practice: the notion of *transparency*, or “clarity of line,” commonly advocated in orchestration treatises, for example, can be understood as a preference for un-clustered layer-level centroids. Clustering relates to what McAdams et al (2022) have termed “textural integration,” where rhythmically independent lines nevertheless appear to integrate into a single auditory stream. These ideas will be explored below in the context of *Music for Small Orchestra*.

Figure 3-9a shows the concurrent layer graphs of Textural Unit 6 (mm. 22–28) and Textural Unit 7 (mm. 28 to 32). Textural Unit 6 appears to cluster into three layers, while Textural Unit 7 clusters into two.[12] We can formalize this clustering by re-appropriating the partitional notation: while rhythmic partitioning tells us how a set of elements is grouped into layers, centroid clustering tells us how those layers are grouped into clusters. We need only redefine the terms of the partitional notation—$x(a, b, c \ldots)$—such that x is equal to the total number of layers, and the integer parts $(a, b, c \ \ldots)$ each represent a cluster of an indicated size. Textural Unit 6 is thus expressed as 7(4,2,1), while Textural Unit 7 is 8(5,3).

[12] As in Chapter 2, the number and size of clusters in these textures were visually estimated and subsequently verified through computational analysis using *k*-means clustering. In cases of more intricate textures where visual estimation of clustering proves challenging, an alternative “silhouetting” approach can be employed to identify the optimal number of clusters (see Rousseeuw, 1987).

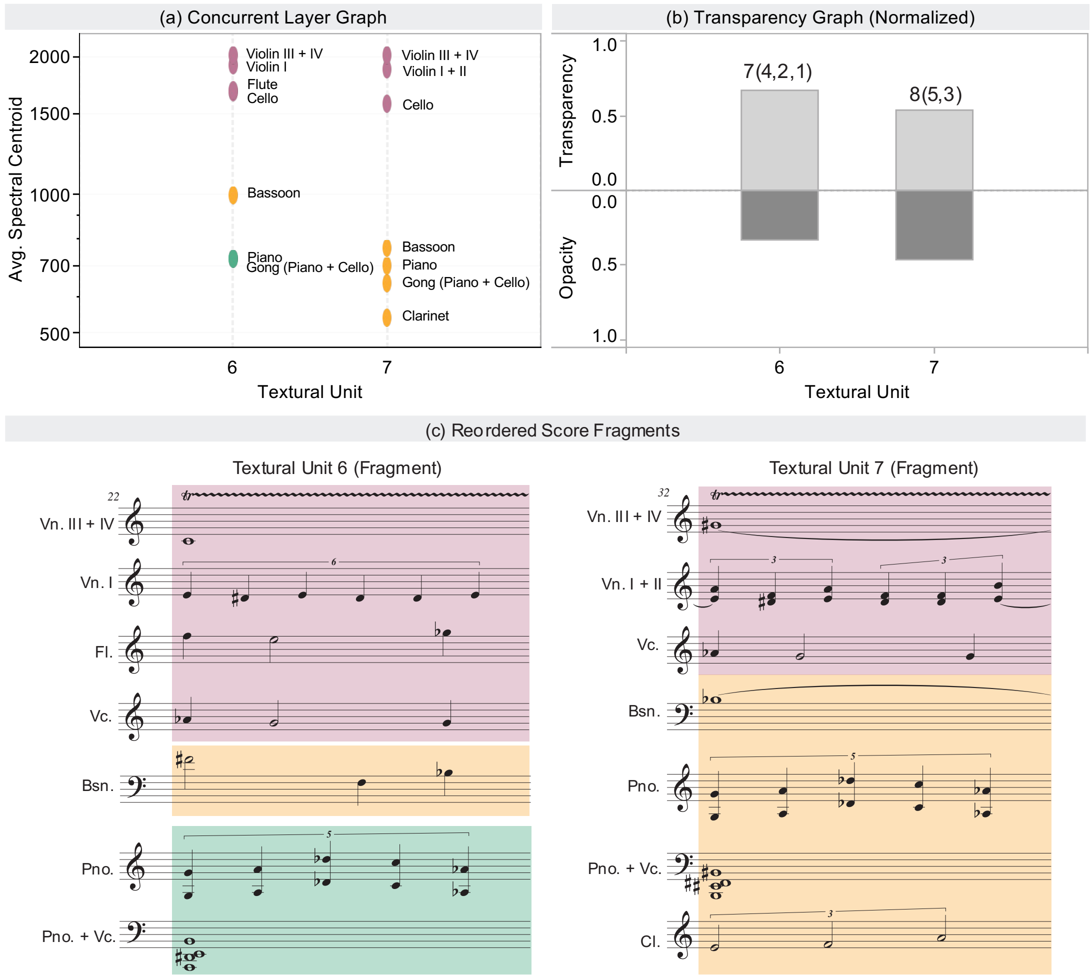


Figure 3-9. (a) Textural Unit 6 groups into three clusters, while Textural Unit 7 groups into two. (b) Textural Unit 7 exhibits greater opacity and less transparency than Textural Unit 6. (c) Score fragments of each Textural Unit reordered according to brightness.

We can then follow the partitional logic further and use these higher-level partitions to calculate agglomeration-dispersion indices, which in this context will not represent coordination but rather the general amount of pairwise layer-clustering, or what I call "transparency." Since the clustering of layers into groups is determined by the magnitude of the SCQs between those layers, a high agglomeration index at this level implies that many pairs of layers are separated by low SCQs, resulting in a high degree of *opacity,* or *muddiness*. A high dispersion index means that many pairs of layers are separated by high SCQs, indicating a clearer, more *transparent* texture. I normalize these values by dividing both indices by the total number of layer-pairs in the texture, which gives a value that reflects the *percentage* of both transparent and opaque pairs.[13] I plot these on a mirrored bar graph, which I call a *transparency graph,* allowing us to visualize the degree to which layers are heard clearly. To illustrate, Figure 3-9b shows the normalized transparency graphs the two textural units in question (6 and 7). As the second of these textures contains more layers, grouped into fewer clusters, the overall texture increases in opacity and reduces in transparency.

Combining concurrent layer graphs with transparency data gives us a useful way of visualizing a larger chunk of music. Figure 3-10 shows the first 44 measures of the piece presented in this way, with the brightness data in the top panel aligned with the normalized transparency indexes in the middle panel. This gives us a general sense of the overall brightness trajectory of the entire passage, as well as granular detail about centroid relationships between layers. The bottom panel isolates a single metric of interest: in this passage, it seems particularly useful to visualize the increasing *opacity* of the texture over time—as layers proliferate and remain unintegrated, layer-clusters grow larger, resulting in

[13] I do this to correct for the fact that textures with fewer layers will always have lower indexes, since they have fewer layer-pairs overall. Expressing the scores as a percentage instead of a count thus squares with our intuition that textures with a small number of distantly related layers are maximally transparent (i.e., 100% transparency).

a more opaque texture overall. Figure 3-10 should be viewed in conjunction with the Appendix, which shows the score of this passage, annotated to correspond with this graph.

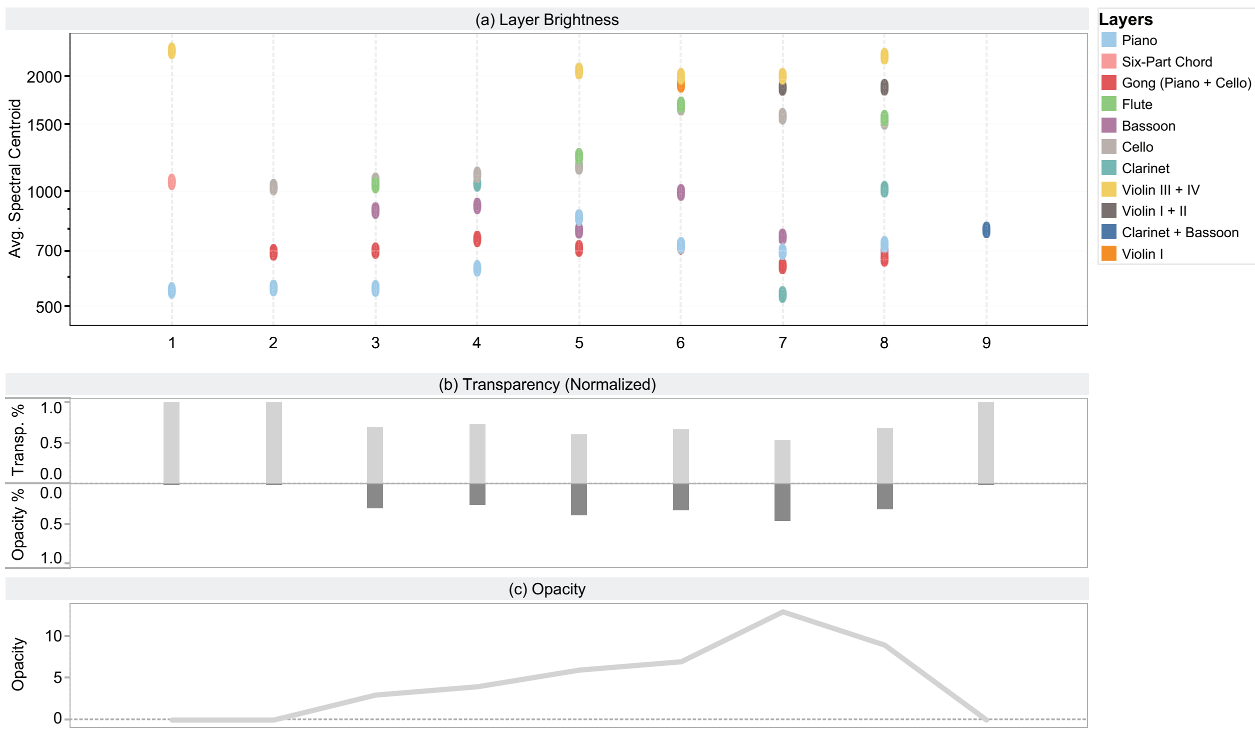


Figure 3-10. (a) Concurrent Layer graph of mm. 3-44 of *Music for Small Orchestra*; (b) Normalized Transparency graph of the same passage, showing fluctuations in transparency and opacity over time; (c) Increasing opacity correlated with more layers and fewer clusters. See Appendix for annotated score corresponding to this graph.

This analysis tells a different story about the piece than does the coordination graph shown above in Figure 3-5. Focusing solely on rhythmic coordination in the piece allowed us to treat the proliferation of rhythmically independent layers (i.e., dispersion) in basically one way: as indicative of increasing perceptual *complexity*. But attending to the relative brightness of these layers paints a more nuanced picture. When new layers are introduced in *Music for Small Orchestra,* they can complicate the texture in at least two ways: on the one hand, they can cluster with other layers, making those layers more obscure. Or they can stake out their own space, increasing complexity by adding to the total number of distinct layers to attend to.

Intralayer Brightness

Measuring intralayer brightness requires us to drill down to the element level once again, to view how the timbre of individual subparts affects the internal homogeneity of a part. While a thorough examination of intralayer relationships in the piece is beyond the scope of this chapter, it will be sufficient to illustrate how this approach complements the interlayer analysis described above: by incorporating element-level detail, we can uncover how layers might overlap—or otherwise appear to be perceptually adjacent—in ways not apparent from concurrent layer or transparency graphs. The discussion here thus mirrors that of Chapter 2, wherein the quantification of the “color” of the “Farben” chords was achieved by examining the centroid distribution of their constituent elements. This approach differs in that results are presented in the context of the complete texture, ultimately aiming to discern perceptual interactions among concurrent layers.

Let us begin by examining the make-up of the six-part chord from Textural Unit 1. Although the layer's overall brightness, as previously observed, falls precisely between two “poles” (the two violins and piano), the individual components of the layer are spread out over a wide range of the spectrum. The left panel of Figure 3-11 plots the spectral centroid of each individual element in the three layers of Textural Unit 1, with the six-part chord in the center. Square brackets connect pairs of elements playing in unison.[14]

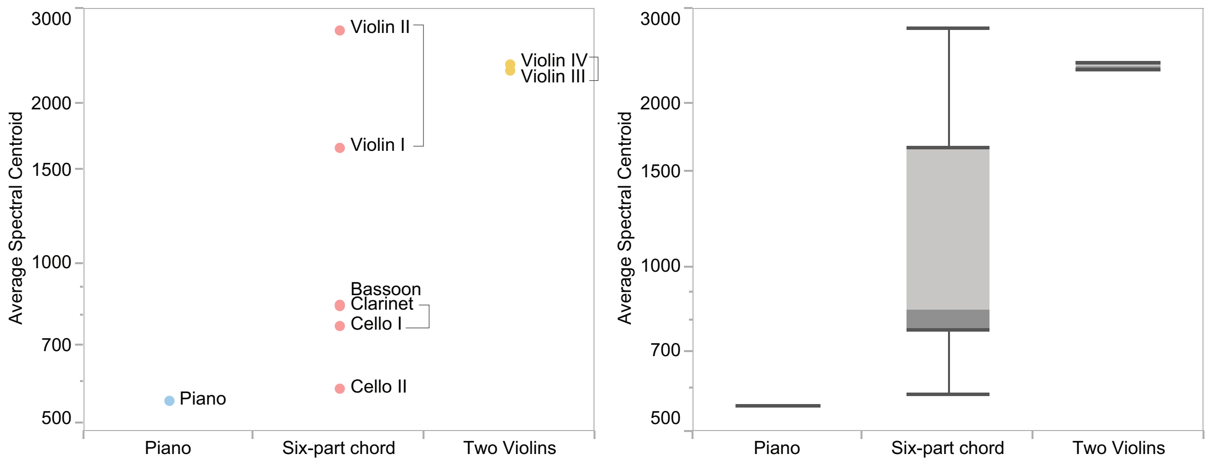


Figure 3-11. Left: Element-level spectral centroids of the layers comprising Textural Unit 1, with unisons connected by square brackets. The six-part chord spans a high SCQ with elements concentrated in the center. Right: Box and whisker plots summarizing the spread of the six-part chord in its textural context.

The previously discussed role of the six-part chord—that it "bridges a gap" between the piano and two violins—is made even more explicit in this visualization: its darkest element is related by a very low SCQ to the piano, while its brightest element is closely related to the violins. The layer's internal distribution can be characterized in terms of clustering: while the range between the brightest and darkest timbres is wide, there is a clear bundle of three

[14] This visualization is somewhat misleading, owing to a peculiar feature of the sample library used: in the MuseSounds sample library, there is a marked difference in brightness between "violin I" and "violin II" samples. This divergence seems to be an intentional aesthetic choice by the manufacturers, with the goal of maintaining perceptual heterogeneity when the two violins play in unison. Other sample libraries deal with this issue differently; thus, in the discussion that follows, I will not place any significance on the SCQ that obtains between the violins in the six-part chord.

elements centered around the midpoint of the layer. The wide range of the elements, coupled with distinct clusters in various spectral regions reflects the perceptual heterogeneity of the layer. In the right panel, I've summarized the overall texture using box plots with vertical lines that indicate the overall span of each layer, a horizontal line showing the median, and a box representing the interquartile range of the layer (i.e., indicating how spread out the bulk of the data is). Since the piano is a single data point, only a single horizontal line is shown. These box-and-whisker plots are helpful for gauging the general uniformity within each layer and estimating how layers might intersect or interact with one another.

Figure 3-12 shows another graph plotting two of three layers in Textural Unit 2: the "gong chord" and the cello ostinato, which persist for the majority of the first half of the piece. Not pictured is the piano's repeated F, which was previously plotted in Figure 3-11. Again, square brackets connect points to indicate which instruments are producing the same pitch. In Figure 3-12a, we can immediately observe the role of the cello (doubling the piano's D2 and A2) in brightening the overall "gong" layer. Put another way, were the cello's perfect fifth to be removed, the left hand of the piano would be the darkest sound in the overall texture. Interestingly, too, the pitch D2 (played by both cello and piano) is simultaneously the brightest and darkest element in Textural Unit 2. This layer clearly divides into two groups based on instrumentation—there is a dark piano cluster and a brighter cello cluster—seeming to suggest a rather heterogenous configuration. Nevertheless, unisons connect these clusters, and encase the entire layer, introducing an important degree of cohesion that ultimately undermines the overall perception of diversity in the layer.

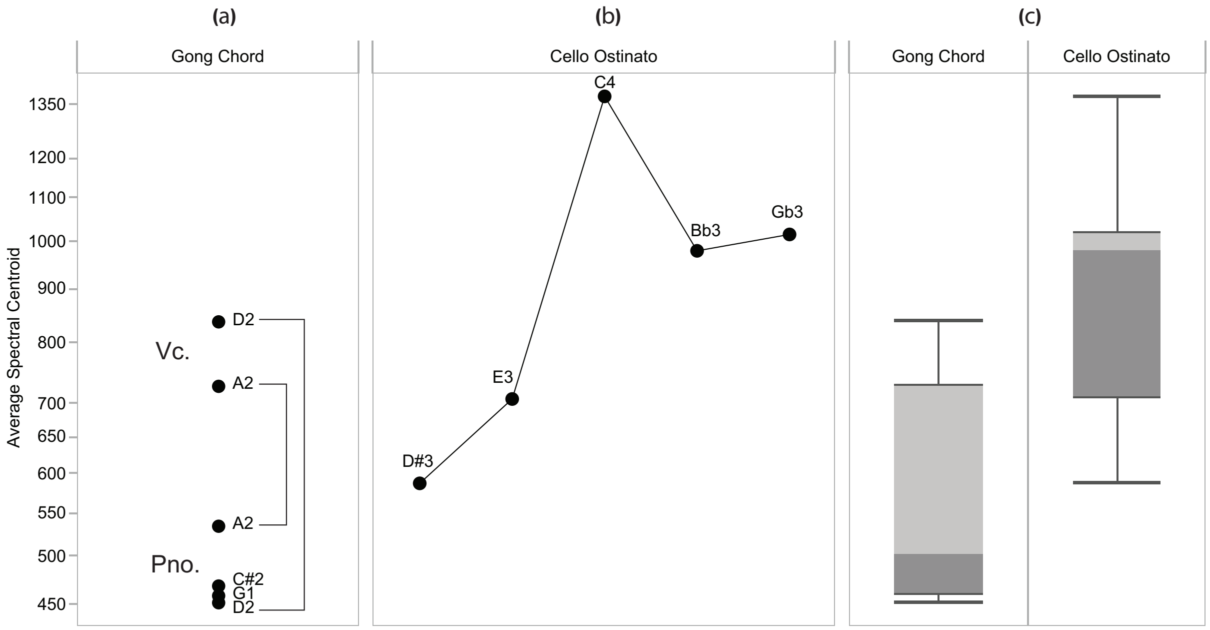


Figure 3-12. Intra- and interlayer brightness relationships of the gong chord and cello ostinato. (a) The “gong chord” contains a bright cello cluster and a dark piano cluster, connected by unisons. (b) The first two notes of the cello ostinato fit into the gap between the gong-chord clusters. (c) Box and whisker plot summarizing both layers.

The middle panel (Figure 3-12b) shows the centroids of the five elements of the cello ostinato, plotted sequentially. Interestingly, the first two notes of the cello fit neatly into the gap between the two clusters of gong chord, while the remaining three notes sit on top of it. There is thus very little “overlap” between these two layers, but instead a sense of interlocking or fitting together in the texture. At the right (Figure 3-12c), the intralayer brightness of both layers is summarized in a box and whisker plot, providing a clear representation of the uniformity and interaction of layers.

Conclusion

This chapter demonstrates the utility of layer-level brightness analysis in revealing the complex interplay of timbre, pitch, and rhythmic coordination in polyphonic post-tonal music. Through an audio-based analysis of interlayer brightness that leverages existing score-based partitional approaches, the dual concepts of clustering and transparency emerged as valuable tools for unraveling layer interactions and tracking perceptual clarity through time. Exploring intralayer brightness through element-level centroids allowed us to visualize the internal dynamics of layers along with additional details about the timbral relationships between them. The findings point to a new development in the analysis of texture and constitute an important building block in a comprehensive multilevel framework. In the next chapter, we broaden our analytical scope further, exploring how analysis of brightness at the level of the *texture* (i.e., the entire audio signal) can reveal further insights that go unseen when focusing on individual elements and layers.

Chapter 4

Texture-Level Brightness and Volatility in Anton Webern's *Six Pieces for Orchestra,* Op. 6/II

This chapter examines the temporal dynamics of brightness across three performances of the second piece of Webern's *Six Pieces for Orchestra* op. 6, in its 1928 revision. My approach employs "texture-level" brightness analysis at multiple temporal scales: overall centroid readings of the complete polyphonic audio signal are averaged over windows of time derived from a preliminary formal analysis of the work. I then relate this brightness-oriented reading of the form to Webern's own understanding of the piece as representing a "catastrophe."

Employing an "always changing mode of expression" (Moldenhauer, 1979), the six pieces of Webern's op. 6 collectively depict a subject's mental state in relation to a tragic event. The first piece, Webern writes in a 1933 program note, conveys the "expectation of a catastrophe," with the second piece depicting the "certainty of its fulfilment."[1] While his published description avoids naming a specific "catastrophe," a private letter to Arnold Schoenberg ahead of the 1913 premiere reveals a deeply personal program: the piece revolves around the death and burial of Webern's mother. Describing the second piece, Webern tells Schoenberg: "It was a beautiful day—for a minute I believed quite firmly that

[1] The third piece functions as an introduction to the fourth, which Webern calls a funeral march, and the fifth and sixth pieces serve as an epilogue conveying "remembrance and resignation" (quoted in Moldenhauer, 1979, p. 128).

nothing had happened. Only during the train ride to Carinthia—it was on the afternoon of the same day—did I learn the truth" (Moldenhauer, 1979, p. 126). In this chapter, I'm interested in examining timbre's role in expressing the structure and meaning of this second piece: how does attending to and visualizing timbral brightness in op. 6/II help us hear Webern's tragic realization of "certainty"? To answer this, I begin by defining the formal units of the piece as they appear at several temporal levels. Then, consulting multiple recordings, I analyze how these units' relationships to each other in terms of brightness give voice to these programmatic elements.

Segments, Blocks, and Sections in Op. 6/II

The short piece, just twenty-seven measures long, is organized into 17 distinct "segments" marked by sudden changes in musical character. These character changes depend largely on orchestration (including both instrumentation and registration), rhythmic content and melodic gestures. On my reading, these segments, ranging from a few beats to a few measures in length, exist at the most foreground level of the piece: they are aggregated to form 7 larger units I call "blocks" and 4 larger units I call "sections." Figure 1 shows a schematic diagram of my temporal segmentation of the piece into these multiple fields. To convey a large amount of orchestral music in a small space, I've chosen to summarize the score in the top panel using an instrumentation graph that simply indicates the presence of instruments (in score order) through time (adapted from Dolan, 2013).

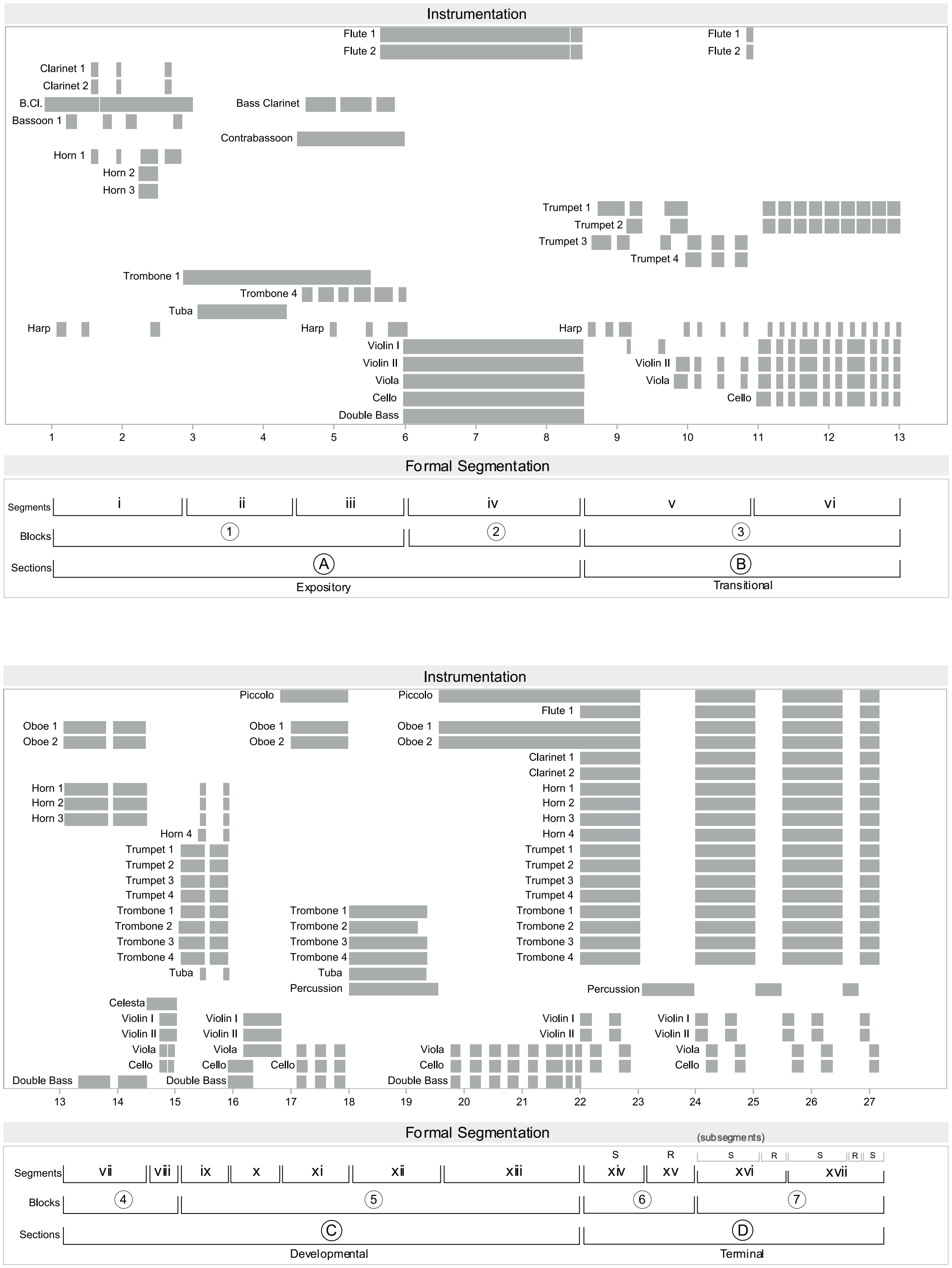


Figure 4-1. Instrumentation graph and formal segmentation of Webern's Op. 6/II. Boundaries at three temporal levels (segment, block, and section) align with changes in orchestration.

As mentioned above, while the 17 segments are a relatively "shallow" formal unit and therefore represent local contrasts, the 7 blocks operate at a deeper level and represent larger chunks of music. Blocks are groups of multiple distinct segments that somehow cohere to form a complete musical utterance or phrase: a sense of continuity persists through the segment-level character changes, and the block as a whole contrasts with the surrounding spans of music.[2] Compared to segments, blocks correlate in much less obvious ways with the surface-level orchestral contrasts (i.e., those immediately visible in the score). The largest unit I define in Figure 4-1 is the "section," which encompasses one or more blocks that again exhibit some form of continuity. At this level, each of the piece's four sections (labeled A–D in Figure 4-1) can be said to perform a broad formal function in the larger structure of the work.[3] The opening section (A) is expository, comprising two blocks, one that I view as introduction-like (1), and one that is more theme-like (2). The second section (B) is a connective transition, building up energy before arriving at a constantly changing—developmental—third section (C). The final section (D) exhibits a

---

[2] This description echoes Hasty's (1984) discussion of post-tonal phrases as "groupings of elements [that] cohere to create a sense of wholeness or completeness" (171). This is similar also to Howland's notion of an "integrated parametric structure" (2015). See also Maler's (2022) listener-oriented study of formal function in post-tonal music.

[3] I coopt four form-functional labels here—expository, transitional, developmental, and terminal—from Temko and Spencer (1994). These are closely related to the notion of formal function described by Schoenberg, Ratz, Caplin, and others. Although these concepts were developed to describe tonal music, I find their generalness useful in a post-tonal context for describing how sections relate to their surroundings (e.g., as beginnings, middles, ends, connections, prefixes, etc.)

terminal function, reaching a dramatic climax that employs the full orchestra, closing the piece with an incessantly repeating pseudo-Mahlerian “scream.”[4]

In the final block (Block 7), I label a handful of small units even “shallower” than segments. These are note-level sub-segments, which can be identified throughout the entire piece, but are labelled only here to clarify the structure of the final block, being as it is a crucial culmination of the piece. Two drastically different sounds make up the final two segments: the aforementioned “scream” (S), played by the full orchestra, and a low “rumble” (R), played by percussion. In Block 6, I read these two sounds (S and R) as distinct segments, but as the rumble element gets shorter, I group them together into a single multi-part segment. In the brightness analysis that follows, I will drill down to this sub-segment level to examine the extreme changes in brightness that occur within these final two segments.

My reading of the form here departs from an apparently similar reading in Baker (1982, pp. 25–26), in that I see no robust connections between non-adjacent blocks or sections in the movement. Baker analyzes the movement as being composed of two unit-types—A and B—which recur in varied form, along with a transition (mm. 9–12) and a codetta (mm. 22–27). For Baker, the primary agent of coherence between non-adjacent iterations of units (e.g., the thing that connects A_1 in mm. 1–5 to A_2 in m. 17) is the network of set classes they share both with each other and with other movements in the work. My reading, which

---

[4] For more on this interpretation of the final chord as a reference to Mahler, see Celestini (2009).

nevertheless aligns with Baker's in several places (see Figure 4-2), places no emphasis on these inter-unit pitch correspondences, emphasizing instead the sonic contrasts that occur primarily as a result of changes in orchestration and register. I find this contrast-oriented reading clearly reflects both Webern's own description of the piece—as involving an "always changing mode of expression" (Moldenhauer, 1979, p. 128)—and a certain volatility of emotion associated with the program. And, while I agree with Baker's use of the label "transition" for mm. 9–12, I refer to the final section (beginning in m. 22) as a "terminal climax" (Osborn, 2013) and view it as the culmination of a formal process, rather than a "codetta" which carries more tonal baggage and undermines the structural significance of the section, tacitly viewing it as extraneous.

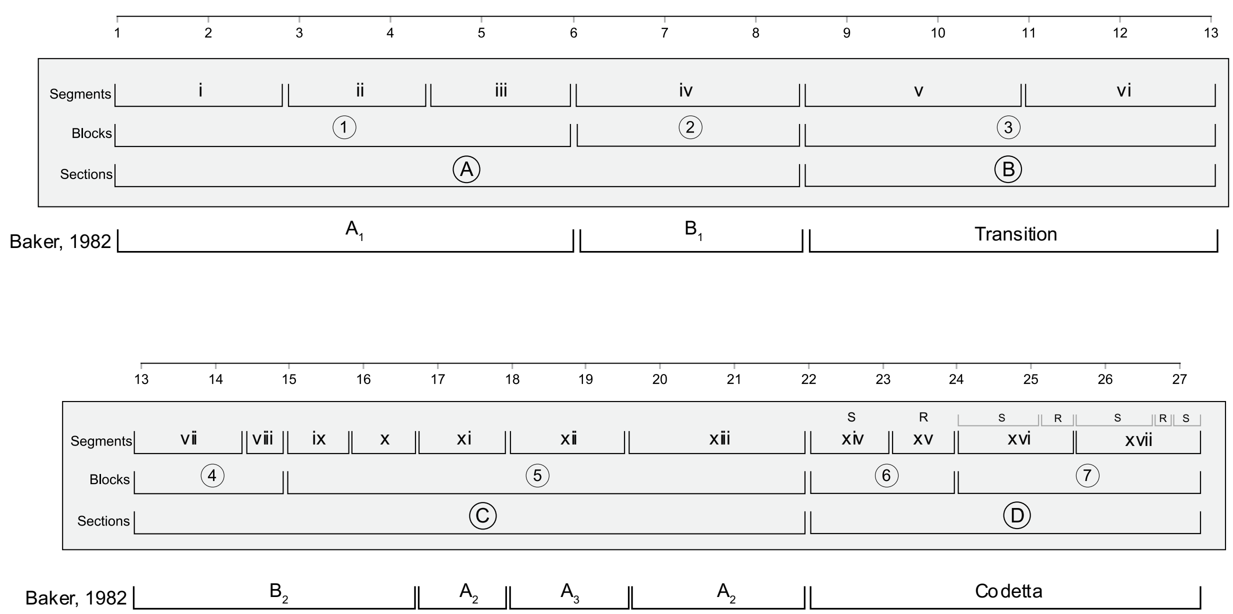


Figure 4-2. My formal segmentation aligned with Baker's (1982) reading of the form of Op. 6/II. Most of Baker's units align with blocks, but some align only with "segments."

Texture-Level Brightness Analysis

While my division of the work into formal units depends on internal coherences and boundary-defining contrasts in multiple domains, I'm interested in characterizing the nature of these coherences and contrasts now solely in terms of timbral brightness. This will allow me to approximate the degree of contrast between formal units and the large-scale timbral shape of the piece. Subsequently, I relate this shape to the expressive concerns of the piece.

Texture-level measurements can be taken at multiple timescales. At one end of the spectrum, for example, a single centroid value can represent the average brightness of an entire piece, while at the other end, thousands of centroids can be strung together to get a picture of the millisecond-to-millisecond fluctuations in brightness. My analysis begins in the middle of these two extremes, taking average centroid readings at two different levels, namely that of the *segment* and *block* as defined in Figure 4-1. Jumping off from my prior formal segmentation, I use an approach I call *sectional averaging:* first, an audio recording of the piece is spliced up in these two different ways (i.e., by segment and block). Then, for each of the resulting audio snippets, the spectral centroid is calculated and averaged over the length of the snippet. This is repeated for multiple recordings of the piece.[5] This

[5] Three recordings are analyzed in this chapter: (1) Robert Craft and Philharmonia Orchestra, 2004 (Naxos 8.557530); (2) Simon Rattle and Birmingham Symphony Orchestra, 2003 (EMI Classics 7243 5 75880 2 7); (3) Herbert von Karajan and Berlin Philharmonic, 1974 (Deutsch Grammophon, 2531 146). These recordings were chosen largely because they feature the 1928 revision, as opposed to the original 1909 version, and are of comparable length and audio quality, thus ensuring a consistent basis for comparison.

technique thus allows me to observe patterns at multiple scales of brightness, and have those patterns be clearly aligned with my reading of the form.[6]

As I described in Chapter 1, a major benefit of analyzing texture-level brightness is that it allows me to engage with actual recordings: issues surrounding source separation and the legitimacy of orchestral sample libraries that arise at other textural levels (i.e., in Chapters 2 and 3) are resolved at this level, opening access to the actual sounds of the work in performance. Salient alignments and discrepancies between performances are thus revealed, allowing us to assess how the decisions of performers produce a wide range of timbral structures that may impact our understanding of the work. Beyond exploring these inter-recording discrepancies, the larger goal of the chapter is to use these performances to delineate an approximate range of possible centroid values brought about by the written composition. In so doing, an understanding of the average brightness of the “work itself” can begin to come into view.

[6] The sectional averaging approach described here depends on an *a priori* formal reading: the division of the piece into formal units occurs before we can analyze the brightness of those units. An alternative approach is to take average centroid readings at regular, form-agnostic time intervals (e.g., every measure, beat, second, audio-frame etc.). This presents its own issues—particularly in terms of temporal alignment of different recordings—but allows us to see contours and patterns of brightness emerge more naturally, in isolation from predetermined formal divisions.

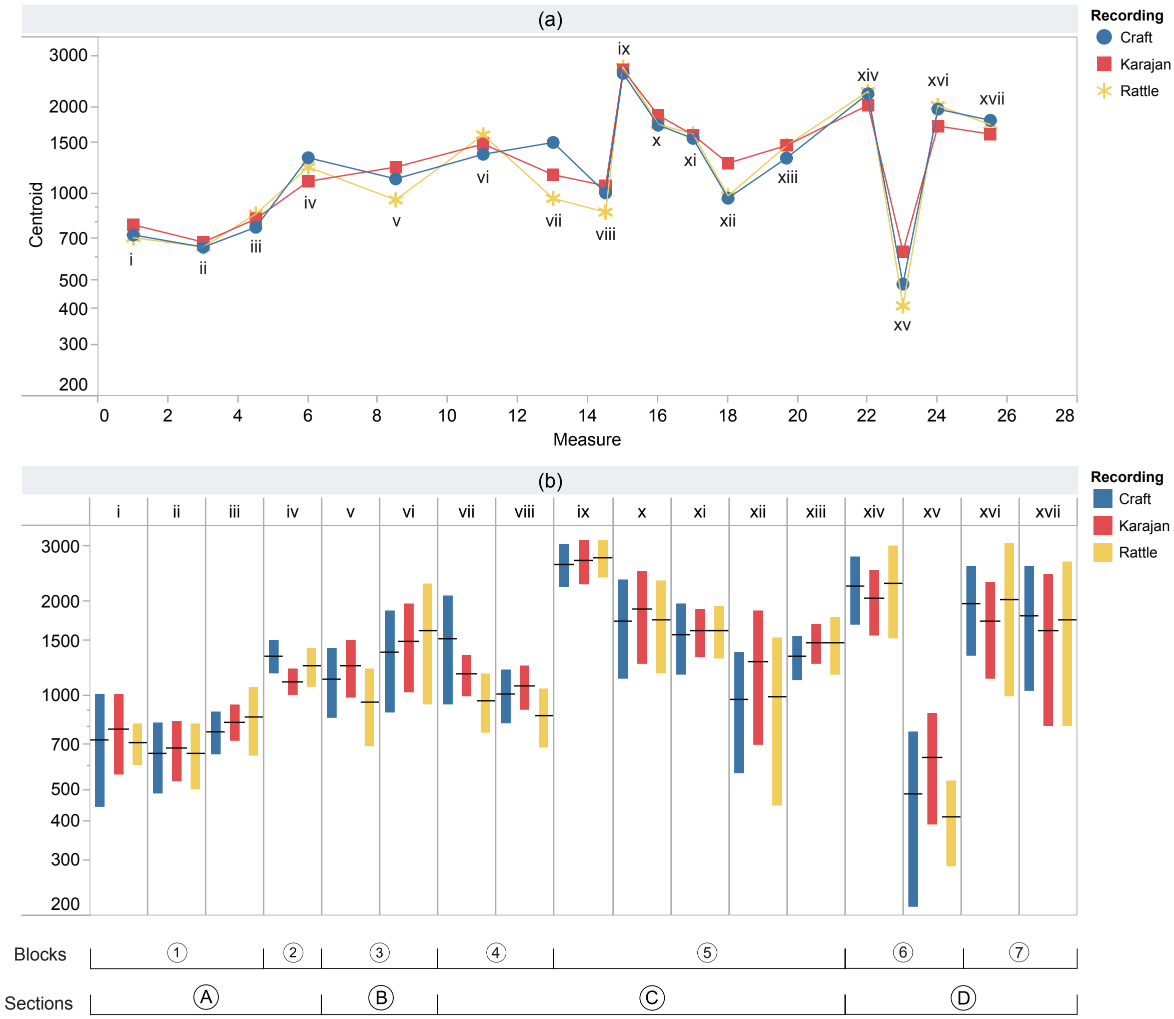


Figure 4-3. (a) Mean centroid ("sectional average") of formal segments of Op. 6/II in three performances. (b) Mean centroid with standard deviation.

Figure 4-3a plots the mean spectral centroid of each formal segment in the three recordings (Craft, 2004; Karajan, 1974; and Rattle, 2003), aligned with the measure numbers in the score. Each data point in 4-3a represents a "sectional average" of a recording of a formal segment. The bottom panel (4-3b) shows the standard deviation for each segment in all three recordings, which represents the degree to which the intra-segment centroids (i.e., hundreds of values at the level of the audio frame) diverge around the sectional average. Beneath the graph, I clarify the grouping of these segments by reproducing the “block” and “section” divisions given previously in Figure 4-1.

At first glance at Figure 4-3, we can detect a great deal of agreement among the three recordings, with their mean centroid trajectories and standard deviations for the most part aligned, and only a handful of noticeable discrepancies (to be discussed in detail below). In all three versions, the first six segments (segments i–vi, blocks 1–3), clearly carry out a gradual brightening of the overall texture. Figure 4-4a below zooms in on these first six segments with a dotted trend line clarifying the steady brightening of the sectional average.

A spike in brightness occurs with the fourth segment (iv), reflecting the change in orchestration that separates Block 1 (harp, clarinets, bassoon, brass, etc.) from Block 2 (flutes and strings). In all three recordings, this timbral change from segment iii to iv is substantial, but the Craft recording exhibits a higher amount of contrast, being both the darkest version of segment iii and the brightest version of segment iv. This is shown in Figure 4-4b, which reports the spectral centroid quotients (SCQ) for each adjacent segment-pair in the first three blocks, highlighting the larger leap (SCQ=1.74) between Craft’s segment iii and iv. Based on aural comparisons of the three recordings, the relative darkness of Craft’s segment iii may be caused by a different type of trombone mute: Karajan and Rattle both feature nasal sounding trombones, while Craft appears to employ bucket mutes, producing a mellower, horn-like timbre. In turn, Craft appears to bring out the bright flute melody in segment iv well beyond the written pianissimo dynamic.

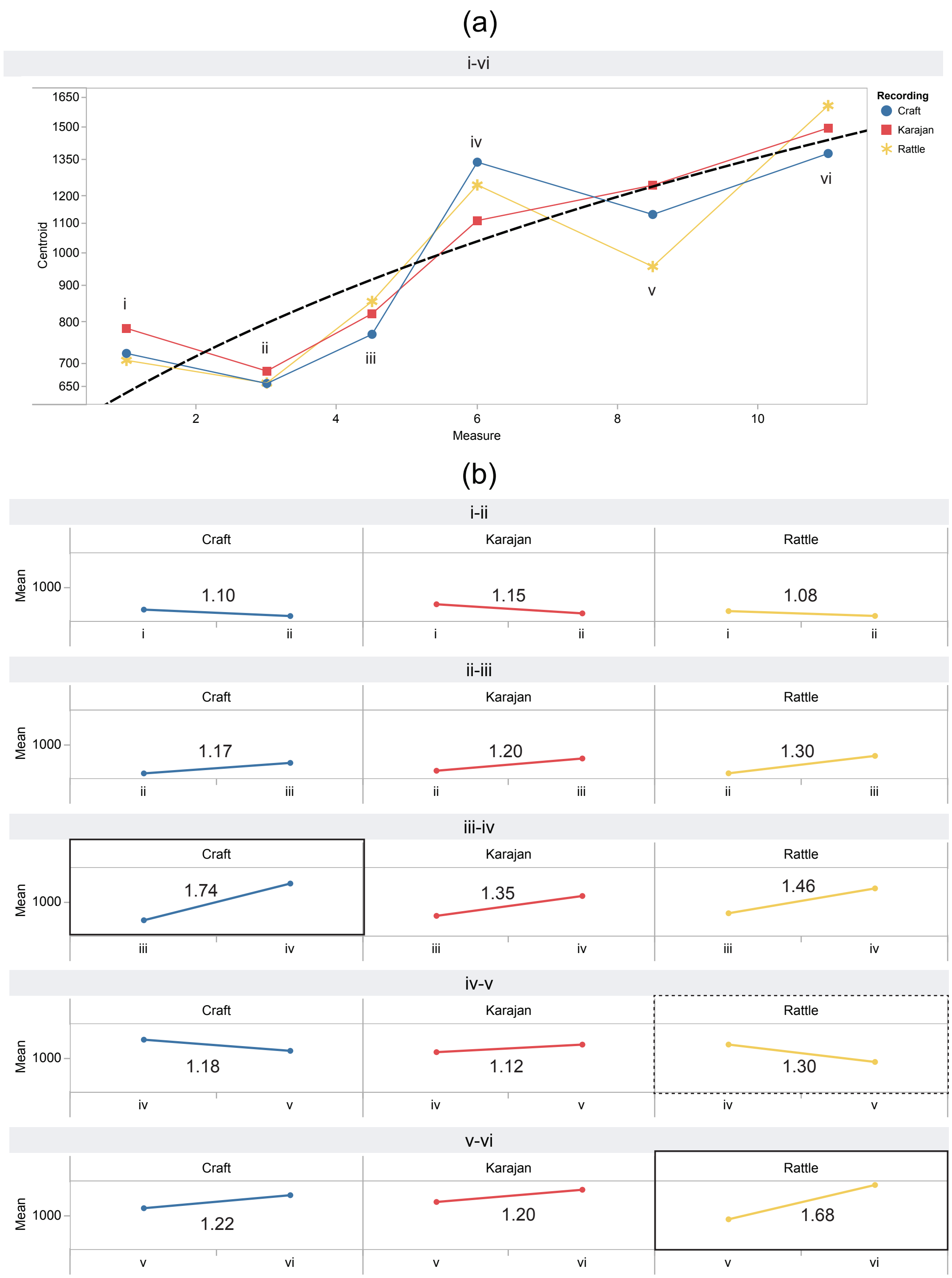


Figure 4-4. (a) Sectional average for segments i – vi across three recordings, showing upward trend in brightness. (b) SCQs for each adjacent segment-pair, outliers shown in bold. Rattle iv – v (dotted rectangle) moves in the opposite direction from the other two.

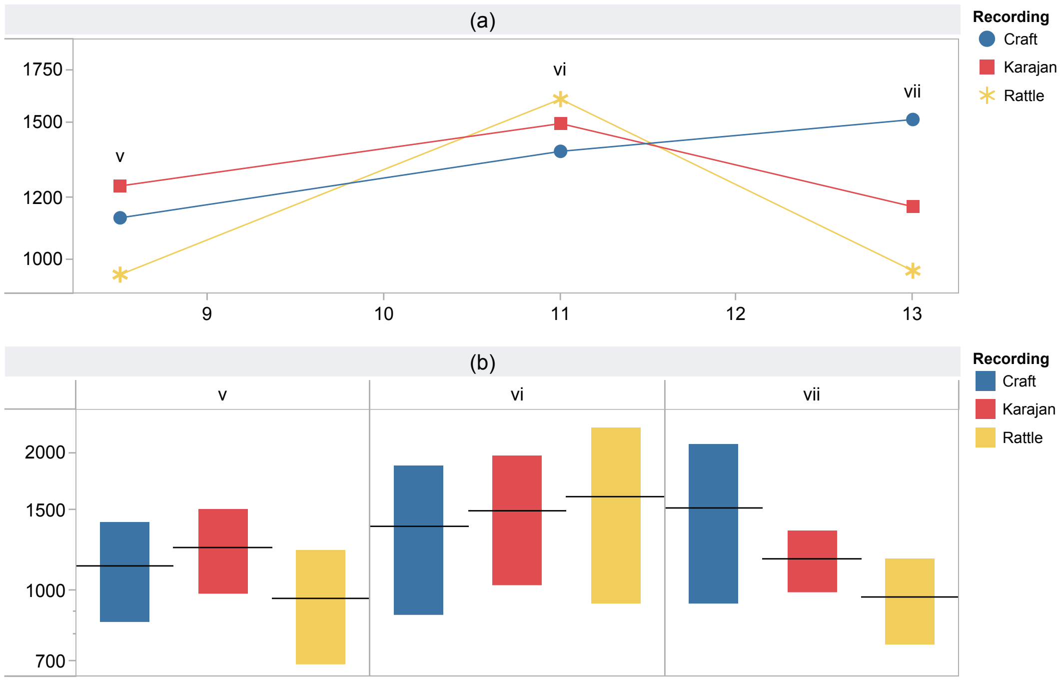


Figure 4-5. Mean centroid (a) and standard deviation (b) for three recordings of Webern Op. 6/II, formal segments v–vii. Rattle's v and vi are more contrasting than the other two recordings (CR = L). Segment vii is widely spaced owing to Craft's "brassier" interpretation.

The fifth and sixth segments (v and vi) comprise the transition section (C), which leads to an arrival at segment vii. Figure 4-5 isolates these three segments. In all three recordings, there is a gradual rise in brightness between segments v and vi, brought about by the rising intensity of the trumpet figure—the growing "brassiness" of the trumpets corresponds to an increase in high-frequency energy. Of the three, the most dynamic interpretation of this transition section (in terms of brightness) is Rattle's: as the steepness of the yellow line in Figure 4-5a suggests, and as Figure 4-4a reveals, the contrast between segments v and vi in Rattle's interpretation is noticeably higher than the other two (SCQ=1.68). In Figure 4-5b, the wider standard deviation bands reflect the crescendo that lasts the duration of segment vi in all instruments (see Figure 4-6 for a reduced score).

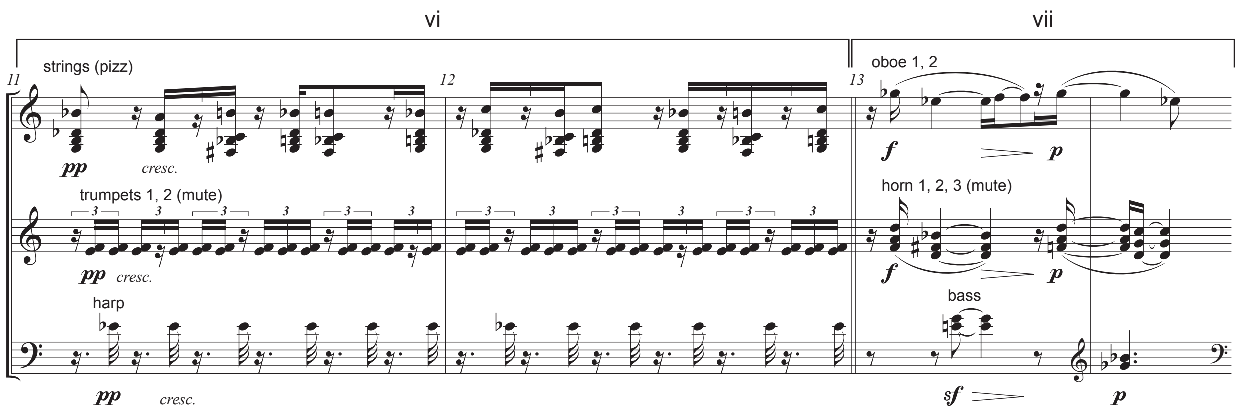


Figure 4-6. Reduced score of segments vi and vii. In segment vi, the crescendo in all instruments in leads to wide deviations in spectral centroid (cf. Figure 4-5b). In segment vii, Craft's interpretation brings out muted horns, while the others bring out the oboe.

The landing point of the transition section is segment vii, which yields the widest range of sectional averages of any single segment (SD=276.06). In Craft's interpretation, this arrival point is *brighter* than the transitional segment (vi) that leads to it (by a low SCQ). This is quite unlike the other interpretations, which show a *decrease* in brightness from segment vi to vii. The reduced score in Figure 4-6 shows that segment vii is scored for two oboes and four muted horns, all at a forte dynamic. In both the Karajan and Rattle recordings, the oboe is clearly the most prominent timbre, with the muted horns assuming a secondary role. Craft's recording, on the other hand, is much louder and brassier, bringing out a piercingly bright horn tone that overpowers the oboe. It seems plausible that the discrepancy here may stem from differing interpretations of the dynamics (*f*) in the score, with Karajan and Rattle viewing them as *performed* dynamics and Craft seeing them as labels of the sonic *result*. In Karajan and Rattle, the horns—playing *forte*—are thus made quieter by their mutes, while in Craft the horns play louder than *forte* to compensate for their mutes.

The resulting timbral difference is of some importance, being as it is a central arrival point in the piece. The brighter, horn-centric version (Craft) connects more clearly with the timbre of the preceding segment: the transition becomes brassier and closes into to a brassy arrival, as if a goal has been attained, or an anticipated event (e.g., a "catastrophe") has occurred. That this segment is a local peak in the Craft version invites us to view it as foreshadowing later such peaks, like those similarly brassy textures at segment ix and xiv. The darker, oboe-centric versions (Rattle and Karajan), on the other hand, sound like a sudden reduction in energy: an implied goal or arrival point, established by the transition section's increase in energy, is now being denied, avoided, or pulled away from.

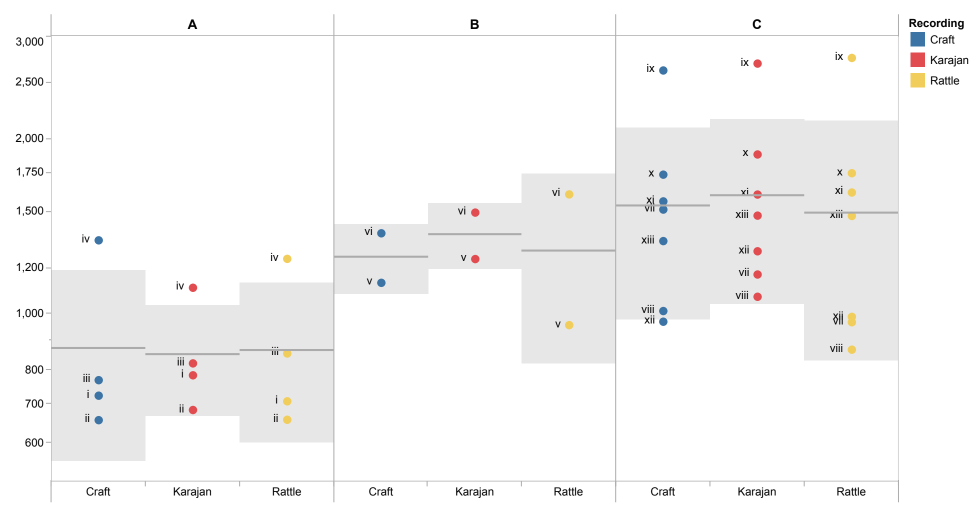


Figure 4-7. Mean centroid of each segment in all three recordings, broken down by section (Section D excluded). Mean and standard deviation given for each section per recording. Section C is on average brighter than the previous two sections.

The remainder of the development section (C) is constantly changing, but on average sits brighter than the opening two sections. This can be seen in Figure 4-7, which compares the segment-level centroids in both Sections A and B with those in Section C and plots the mean centroid of each section. The wider standard deviation bands in Section C (SD = 565.86 Hz) indicate more volatility as compared to Section A and B (SD = 313.93 Hz): individual segments sit at higher peaks and lower valleys, further away from the mean centroid of the section. The section as a whole is excerpted in Figure 4-8, which shows the segment-level mean centroids (4-8a) and standard deviation (4-8b).

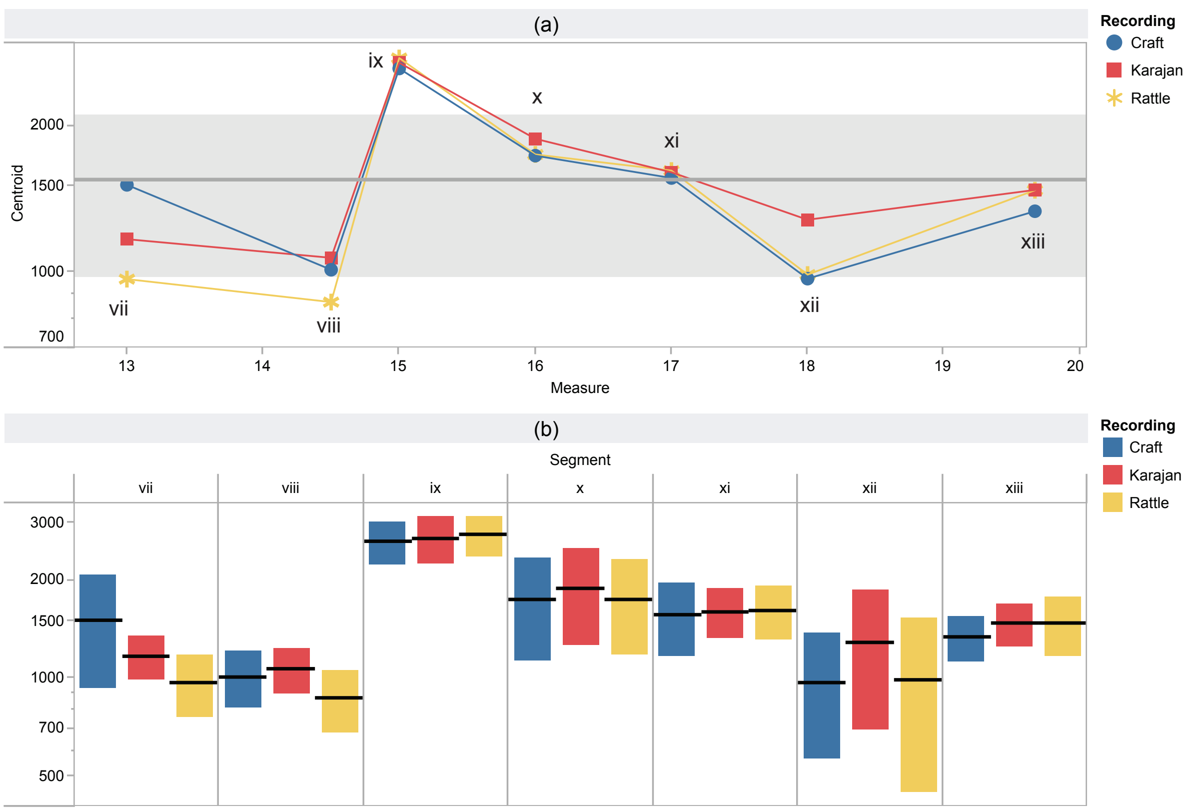


Figure 4-8. Sectional average for segments vii – xiii across three recordings, showing mean and standard deviation of the section. Segment ix is the highest peak of the piece.

A pivotal moment in the development section, and the piece as a whole, is the howling brass tutti—played fortissimo, with mutes—in segment ix, the highest segment-level peak in all three recordings (Avg. SC = 2700 Hz). In all recordings, a high centroid ratio (SCQ) obtains between segment viii—a dark and tender moment for celesta and strings—and segment ix (Avg. SCQ = 2.76). The fact that segment ix reaches the piece's highest peak in brightness prompts us to place special emphasis on its structural significance and invites us to associate it with later peaks in the terminal climax (explored further below). This important moment is followed by a transitional segment (x) in the strings that rapidly traverses a four-octave pitch range and leads to segment xi, a shrieking, piccolo-dominated motive that resembles a warped steam whistle.[7] After a suddenly dark interjection from the trombones and bass drum, foreshadowing the funeral march of the next movement (segment xii), the steam whistle returns (segment xiii) and leads us to the terminal climax (D).[8]

---

[7] The mobility of segment x reveals the necessity of pairing sectional averages with standard deviation—in Figure 4-8a, the single point for segment x is misleading, since it traverses four octaves and yields a wide range of centroid values. This wide range is captured in Figure 4-8b, however, by wider SD bands.

[8] To my knowledge no existing analysis connects this movement to train sounds. While it won't be developed here, it seems to me a promising angle for interpreting the movement's various screams, howls, whistles, and locomotive rhythms: Webern's letter to Schoenberg, after all, discloses that his tragic realization occurred on "the train ride to Carinthia."

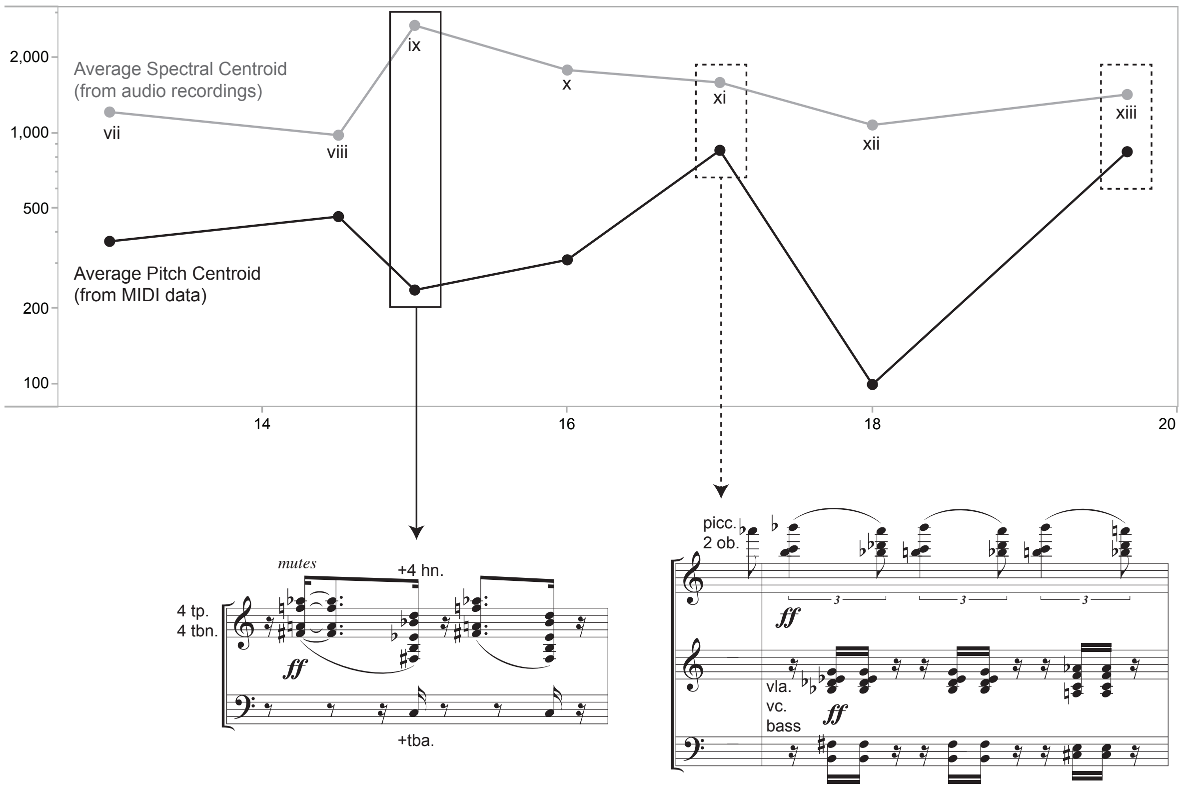


Figure 4-9. Average spectral centroid and pitch centroid of Section C. Dotted rectangles show peaks in pitch (corresponding to the "steam whistle" motive), solid rectangle shows peaks in brightness (muted brass tutti). Peaks in brightness do not coincide with peaks in pitch.

As we've seen in previous chapters, peaks in brightness do not always align with peaks in pitch. Indeed, if we were oriented solely toward pitch height during the development section (C), the steam whistle motive (segments xi and xiii) would be the high points of the section. But taking overall brightness into account reveals that the rich spectrum of the full brass section, playing fortissimo into mutes causes segment ix to outshine the piccolo/string combination in segments xi and xiii, despite its lower average pitch. Figure 4-

9 displays this information, giving the spectral centroid of each segment (averaged across the three recordings) aligned with the segment's average *pitch centroid*.[9]

The climactic closing section is comprised of two opposed components: the bright "scream" (S) and the dark "rumble" (R). S involves, for the first time in the piece, the entire orchestra together: the full brass section, with mutes, is combined with the piccolo and woodwinds in their highest registers, strings producing snap pizzicato, and a loud crash cymbal roll. The forceful combination yields the second highest segment-level peak in the piece (Avg. SC = 2185 Hz).[10] R, on the other hand, is quietly played by a deep bass drum, and represents the lowest (darkest) valley in the piece (Avg. SC = 512 Hz). Figure 4-10 drills down to the level of the sub-segment from mm. 24–27 and labels SCQs between them, revealing the high degree of contrast between S and R that characterizes the final two segments of the work.

---

[9] The pitch centroid is computed by running a centroid calculation on a pure sine tone realization of the MIDI data of the segment, averaged over the length of the segment, in much the same way that the spectral centroid data is collected. Since the spectral centroid of a sine tone is also its fundamental frequency, the spectral centroid of a polyphonic sine-tone realization gives us the center of gravity of all the fundamental frequencies (i.e., the pitch centroid). This method was chosen to keep the sectional averaging approach consistent throughout the chapter and thus facilitate direct comparison.

[10] Here, the texture-level centroid alone fails to capture the enormous weight and unique spacing of this final chord. As I will describe further at the end of the chapter, it is significant that this chord is the endpoint of a large-scale accumulative instrumental process. Element-level analysis may also reveal important form-defining details: it seems very likely, for example, that the piccolo's piercing G7 here is in fact the brightest element in the piece.

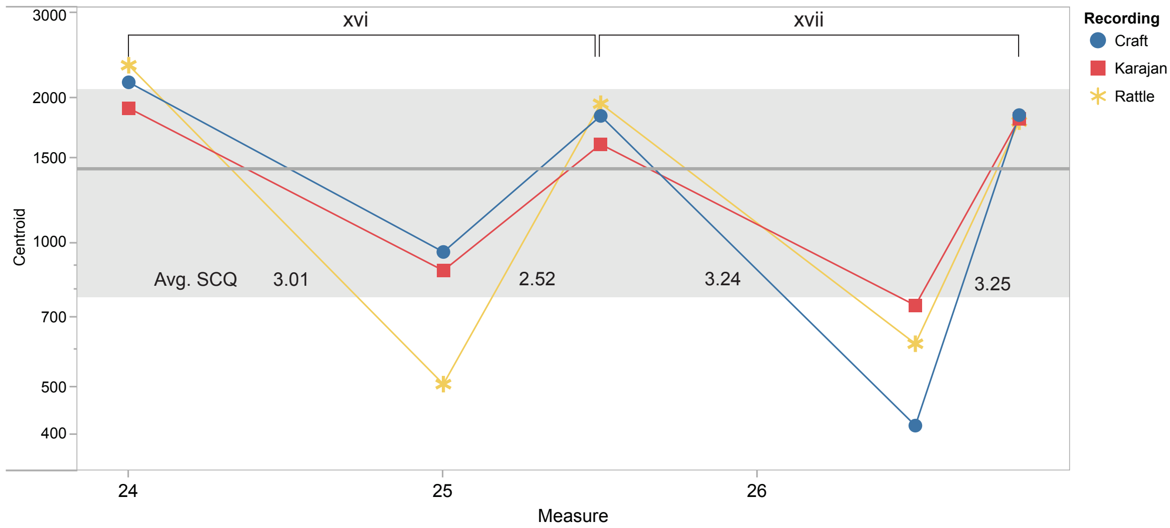


Figure 4-10. Subsegments from mm. 24–27 across three recordings. Average SCQs between subsegments reported.

Over the course of the piece, segment-to-segment changes in brightness become more extreme, with increasingly higher peaks and lower valleys. In segments i through vi, brightness changes between adjacent segments are generally slight (Avg. SCQ = 1.25), despite some medium-contrast spikes (i.e., between iii and iv). As described above, the development section (C) exhibits larger and more frequent contrasts, while the terminal climax (D) delivers the work's most extreme shifts. I call this attribute volatility and quantify it by calculating the standard deviation of the segment-level centroids during the section in question. Thus, higher volatility scores indicate that a section has higher peaks and lower valleys—i.e., larger segment-level deviations (in frequency) from the mean brightness of the section. Another method of showing volatility is to examine the average SCQ of a section, with higher values indicating greater contrast between adjacent segments in the section.

In the foregoing analysis I emphasize two claims: the piece exhibits an upward trend both in terms of (1) its overall brightness and (2) its volatility. This is summarized in Figure 4-11. The top panel (4-11a) shows the average brightness of the three recordings at the segment level, with bands indicating the spread of the three recordings, and a dotted line illustrating the upward trend in brightness. The bottom panel (4-11b) divides the same data up into four sections, with the mean and standard deviation given for each section. The average SCQ for each section is reported in Figure 4-11b.

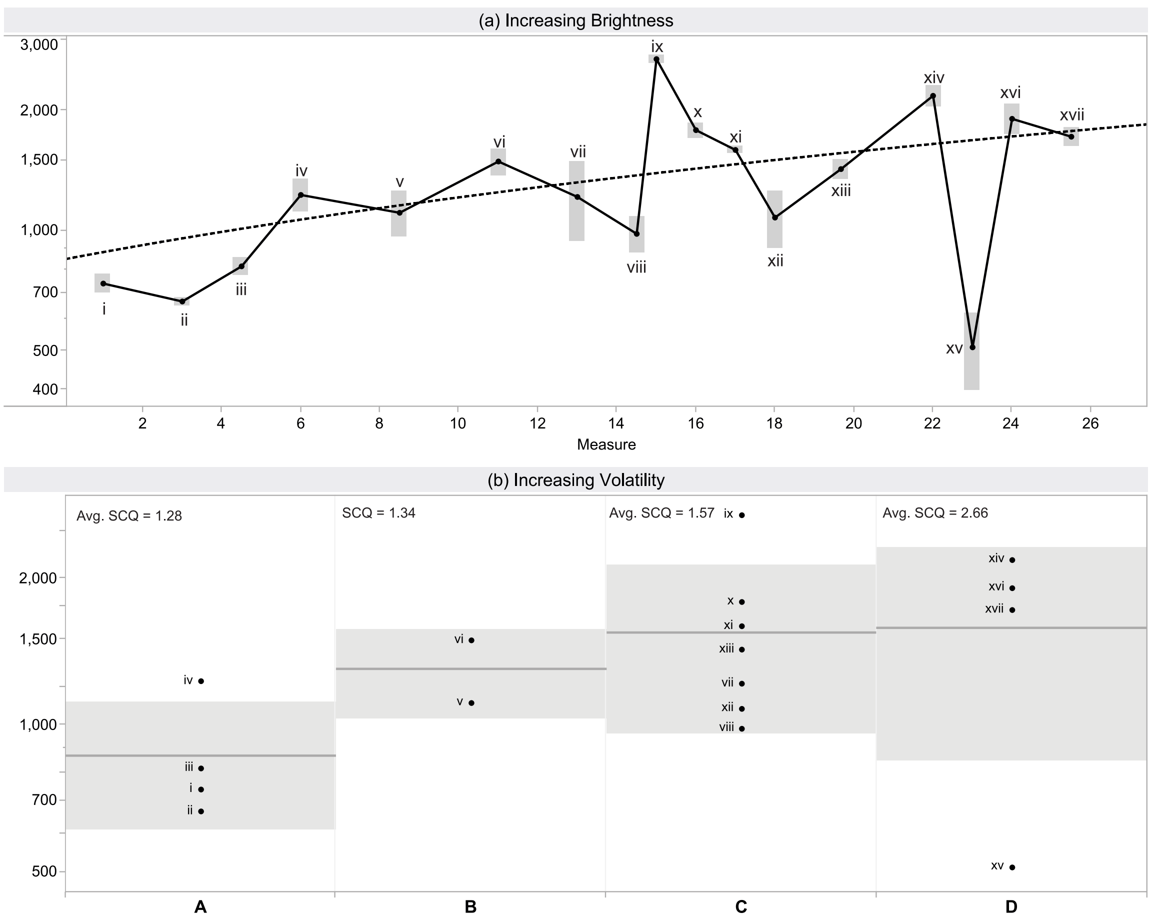


Figure 4-11. (a) Segment centroids averaged across all three recordings. Bands show the spread of recordings. Dotted line shows increasing brightness trend. (b) Segment centroids broken down by section. Standard deviation bands widen with each section, and average SCQs rise, reflecting increased volatility.

There are many conceivable ways of relating this analysis to the emotional and narrative dimensions of the work; here I will offer one reading. I interpret the two opposing poles of the piece—brightness and darkness—as representing an opposition between acceptance and denial of Webern's tragic event. The piece thus enacts a gradual process of "taking in" or comprehending the catastrophe. Moments of abruptly reduced brightness suggest denial, resistance, or a struggle to acknowledge the magnitude of the tragedy, while peaks in brightness suggest clarity, certainty, acceptance, and the pain that goes along with them. The relative darkness of the opening three segments depicts an initial state of unawareness, as if "believing quite firmly that nothing had happened." Segment iv, suddenly brighter, jolts us awake to the possibility of a catastrophe whose "certainty" is revealed later in segment ix (the brightest peak of the piece). The lead-up to this moment—the path from segment iv to ix—is characterized by undulation: first darkening to segment v, then brightening through the transition (vi), and darkening again at the onset of segment vii, like a brief wave of lucidity stifled. The dark tenderness of segment viii—played by low celesta and strings—represents the nadir of this process, providing a fleeting moment of calm before abruptly giving way to segment ix, which I view as the crucial moment of illumination. The ensuing development—segments x to xiii—resembles the previous "wave" in segments iv–ix, but now reversed: a downward spiral of denial, arriving at a low point in segment xii, is counterbalanced by a brightening motion that leads us to the full-orchestral "scream" of segment xiv. Here we find the most extreme and frequent alternations between acceptance and denial, as Webern grapples vividly with the fact of his mother's death. That the bright peak of the "scream" motive ultimately wins out over low rumbles of intense denial and becomes the final sound of the movement suggests his final submission to a painful truth.

On this reading, there is a programmatic connection between segments iv, ix, and xiv, all of which are bright points representing realization. Examining the instrumentation at each of these peaks, we find a gradual accumulation, with segment iv scored for two flutes and strings, segment ix for the full brass section alongside low strings, and segment xiv for the entire orchestra. Each peak is scored for more instruments, gradually filling in those instruments missing at the previous peak, leading to the massive full-orchestral chord that closes the work. This accumulative process can be interpreted as the tragic sense of certainty taking over, saturating the mind and body until it can no longer be denied. This reading is supported by the fact that dark points in the piece—moments of denial at segments v, viii, xii—gradually decrease in instrumental numbers, until we are left with only a rumbling bass drum (segment xv), a sole remnant of resistance.

Conclusion

This analysis shows that visualizing statistical trends in overall brightness and volatility provides a promising foundation for interpreting the formal and emotional dimensions of Webern's Op. 6. An important direction for this analysis would be to examine more closely the pitch language of the work. This could begin by exploring its evolving "pitch centroid" (shown previously in Figure 4-9) and proceed through a deeper analysis of the harmonic makeup of segments, blocks, and sections. Another direction is to use the analytical strategy to compare Webern's multiple orchestrations of the piece. Here we examined only the “definitive” 1928 revision of the work, but, as Falck (1993) reviews, there are several important differences in the 1909 version that might impact a timbral analysis. Webern also arranged the work for a small chamber ensemble; one might thus examine how this reorchestration preserves or disrupts the brightness trajectories revealed in this chapter.

More generally, this chapter demonstrates the analytical utility of analyzing brightness at the level of an overall texture, the highest level defined in Chapter 1. By scrutinizing statistical tendencies of the overall signal, in multiple realizations and recordings, we gain insight into trends and patterns that might otherwise be obscured by the particulars of individual layers and elements.

Chapters 2 to 4 have collectively demonstrated the capacity of each textural level to stand as the focus of an in-depth analysis. Now, these levels can be alternated, integrated, or otherwise combined in a comprehensive multilevel timbral analysis of a complete piece. This integrative approach will be the focal point of the final chapter.

Chapter 5

Timbre as an Agent of Delineation in Edgard Varèse's *Octandre,* I.

Despite Edgard Varèse's abundant remarks on the role of timbre in his music, relatively little of his output has been analyzed closely for its timbral structure. Analytical scholarship on Varèse's music has concentrated on pitch, while timbre-oriented analysis has remained either superficial or neglected altogether.[1] This chapter responds to this neglect by presenting a multilevel analysis of timbral brightness in the first movement of Varèse's *Octandre* (1923) for seven wind instruments and double bass. I begin with a discussion of timbre in Varèse's music, with reference to both the composer's own words and the secondary literature. Then, I proceed through a close reading of the first movement, integrating the various levels of brightness discussed in the previous three chapters (*element*, *layer*, and *texture*) and considering how these levels contribute to the articulation of the movement's formal structure.

Varèse on Timbre

Varèse often spoke in general terms about his compositional practice, using an enigmatic lexicon that eschewed conventional musical terminology. "Planes," "surfaces" and three-dimensional "masses" of sound are consistent analogies, evidently central to his thinking but never applied to analytical commentary on specific passages or pieces. These

---

[1] Two major publications on timbre in Varèse are worth mentioning as exceptions here: Francois (1991) and Chou Wen-Chung (1978). Both focus on *Ionisation* (1929–31) and adopt an instrument-oriented approach to timbre analysis.

analogies, thus, offer limited practical insight into his methods: determining exactly what a "sound mass" is—and where to find one—has been discussed extensively in the analytical literature, but appears nevertheless to remain an unanswered question (Chou, 1966; Erickson, 1975; Strawn, 1978; Anderson, 1984, 1991; Smoot, 1986; Bernard, 1987; Reynolds, 2013).

Consequently, scholarly opinions are divided regarding the extent to which Varèse's analogies should guide the analysis of his music. Erickson (1975) insists that Varèse's descriptions of sound masses are "serious and precise," and the phenomena can be identified in his works in the form of "fused ensemble timbres." Anderson (1984, 1991), Strawn (1978), and Smoot (1986) each locate specific examples of Varèse's sound masses resembling the "fused ensemble timbres" found by Erickson.[2] Bernard (1987) interprets Varèse's concepts somewhat more abstractly, translating his notion of "space" into pitch space (in semitones) plotted against time. Rather than pin down exact occurrences of Varèsian concepts, Bernard constructs a "pitch/registral" theory that on the one hand loosely aligns with Varèse's thinking—e.g., emphasizing the vertical dimension (and its partitioning) and processes of continual transformation—and on the other hand centers theoretical concepts unmentioned by Varèse (e.g., pitch symmetry). Moura (2004), on the far end of the spectrum, cites scholarly disagreement as justification for setting aside the

---

[2] Chou Wen-Chung (1966) discusses sound mass without reference to "fusion." Differences between Anderson's approach and that of Chou are covered in Moura (2004). Erickson considers "fused ensemble timbres" to be a kind of "sound mass," but lacks a detailed account of what qualifies as "fused" (see Swift, 1975).

composer's words altogether and expanding on certain symmetrical concepts explored in Bernard.

Of particular significance for the current project are Varèse's published comments on timbre, which have brought about similarly divided opinions. Lecturing in 1936, Varèse famously said that timbre (in electronic music) would become "an agent of delineation like the different colors on a map separating different areas, and an integral part of form" (Varèse & Chou, 1966, p. 12). He was intrigued by certain "acoustical arrangements" that would render his masses, planes, and beams of sound clearly audible and distinct from each other, achieving a "hitherto unobtainable non-blending" (12). Varèse refers here to the potential of electronic music, but his statements about orchestration closely mirror these prior musings, revealing attempts to create with instrumental timbre these sensations of "non-blending":

> Orchestration is an essential part of the structure of a work. Timbres and their combinations—or better, quality of tones and tone compounds of different pitch, instead of being incidental become part of the form, coloring and making discernible different planes and sound-masses, and so create the sensation of non-blending. Variations in the intensity of certain tones of the compounds modify the structure of the masses and planes (quoted in Reynolds, 2013, p. 223)

In both quotations, Varèse's cites the familiar distinction between that which is "structural" or "integral," and that which is merely "incidental." While pitch and rhythm are typically taken to be the primary carriers of compositional logic, timbre is often understood as a surface feature and reduced out in analysis (without threatening, say, to render the analyzed piece unrecognizable or incomprehensible). Varèse, however, wants to claim for timbre a much more salient role as structural determinant, or carrier of form: for the composition to unfold according to Varèse's design—i.e., for a sensation to be induced of unrelated sound masses growing, moving, and interacting in space—concurrent sonic layers must each be cast in a unique color that renders them perceptually distinct from each other.

Interpreting Varèse's statement above—that timbre is an "agent of delineation"—Bernard (1987) asserts that "for Varèse, timbre is a partitioning device" (46): what is being made "discernible" by timbre is ultimately a collection of fundamental frequencies, whose organization determines the integral structure of a composition. Timbre, on this reading, is delineative in the sense that it clarifies boundaries between pitch structures, disambiguating cases where layers overlap registrally. Guck (1992, 1993) takes issue with Bernard's pitch-centric approach, which denies a structural role for timbre and dismisses Varèse's understanding his own music. Bernard's defense (1992) agrees that timbre is indeed "elevated" in Varèse's music but argues that the inherent resistance of instrumental timbre to precise "control" renders both systematic analysis and methodical composition of timbral structures unfeasible.[3]

[3] Bernard also takes issue with approximations and appeals to individual subjectivity, which he sees as endemic to timbre-focused analysis. He cites Erickson (1975) and Strawn (1978) as examples of a kind of vagueness that permeates the scholarship on Varèse.

This dissertation has thus far shown that instrumental timbre is controlled via the selection and combination of instrumental sounds that have predictable sonic characteristics desired by the composer. While composers can only pin down a *range* of possible realizations in sound, that does not make their decisions unsystematic, nor timbral analysis vulnerable to excessive appeals to subjectivity, as Bernard worries. There is no reason to believe that Varèse was not thinking systematically about the palette of sounds available to him, nor able to exact enough "control" over it to warrant close analysis.[4] And, as this dissertation has demonstrated, such close timbral analysis is indeed possible, and can be both grounded in empirical measurement and combined with a pitch-focused reading. Thus, while the multilevel analysis that follows will center timbral brightness, it leaves open intentionally the possibility of integrating pitch-based insights from existing analyses.[5]

[4] Although he didn't use the word "control," Varèse evidently believed composers should understand the precise effects of instrumental timbre, writing that it is "not until the air between the listener's ear and the instrument has been disturbed does music occur.... In order to anticipate the result, a composer must understand the mechanics of the instruments and must know just as much as possible about acoustics" (quoted in Strawn, 1978). Lalitte (2011) has demonstrated the role that acoustic theory, namely that of Helmholtz, played in shaping Varèse's theory and practice of composition.

[5] Pitch-based readings of large portions of *Octandre* are found in Bernard (1987), Morris (1997), Anderson (1984), Moura (2004), Lalitte (2011). Di Gasbarro (2018) provides a useful discussion of Varèse's conception of chromatic pitch space—emphasizing wide vertical spacings of interval class 1 (e.g., major sevenths and minor ninths)—with reference to Varèse's pitch diagrams found in the Sacher collection (see also Chou, 2006).

Levels of Timbral Design

Evident in comments by Varèse and others is the notion that timbre operates at multiple scales of detail. In his argument for the "essential" role of timbre quoted above, for example, Varèse seems to bring out three subtly distinct aspects: (a) timbre contributes to a plane or mass' individual "color," (b) that color allows the plane/mass to be distinguished from (or "non-blended" with) other simultaneous sounds, and (c) through this process, timbre becomes an "integral" part of form. Each of these ways of framing timbre—as the quality of a verticality, as the thing that distinguishes simultaneities, and as a structuring force—is related to a level of brightness analysis described in this dissertation: the level of the *element*, *layer,* and *texture* respectively. Each of these will be expanded upon in turn.

To get at the color of a melodic line or verticality we can examine the individual components that comprise it via element-level analysis. Both Cowell (1928) and Babbitt (1966) point to this way of understanding Varèse's music. Cowell explains: "besides the harmony of notes, which, with Varèse, is somewhat secondary, there is at any given time also a harmony of tone-qualities." While certain of his pitch collections "can be found in many a modern composer's work," their scoring brings about a "harmony" resulting from the timbres of the instruments in their particular registers that is distinctly Varèsian. Echoing this, Babbitt writes: "the effect of different 'harmonies' is by no means dependent entirely on the explicit pitches presented by each instrument, but most importantly, on the strikingly different spectra associated with these instruments ... as a result of the different registral placement of the fundamental in each instrument" (1966). By examining the spectral centroid of individual elements, along with the centroid-distribution of a sonority (or sound mass) we can be more specific about how this "harmony of tone-colors" works—how the spectral properties of individual elements combine to produce a colored mass of sound.

Further, to analyze how "acoustical arrangements" bring about the "non-blending" of unrelated sound masses, we can estimate perceptual distances between simultaneous layers as we did in Chapter 3, comparing layer-level centroids of groups of rhythmically coordinated elements. Whereas element-level analysis gives us a way to talk about individual colors, layer-level analysis gives us a way to talk about the textural effect of those colors, i.e., the extent to which they bring about the perceptual partitioning of the music into distinct layers.

Finally, timbre's "integral" role in form, while it depends on these lower levels, can further be analyzed through a texture-level approach, revealing trends and trajectories of the entire polyphonic totality. This relates to what Babbitt (1966) calls the "total spectrum" in Varèse's music. In a related vein, Cogan's (1984) spectrographic analysis of Varèse's *Hyperprism* focuses on the complete audio signal (i.e., the texture level), but my approach at this level relies more on statistical trends of the overall centroid, averaged at multiple timescales, than on the information visible in a spectrogram.

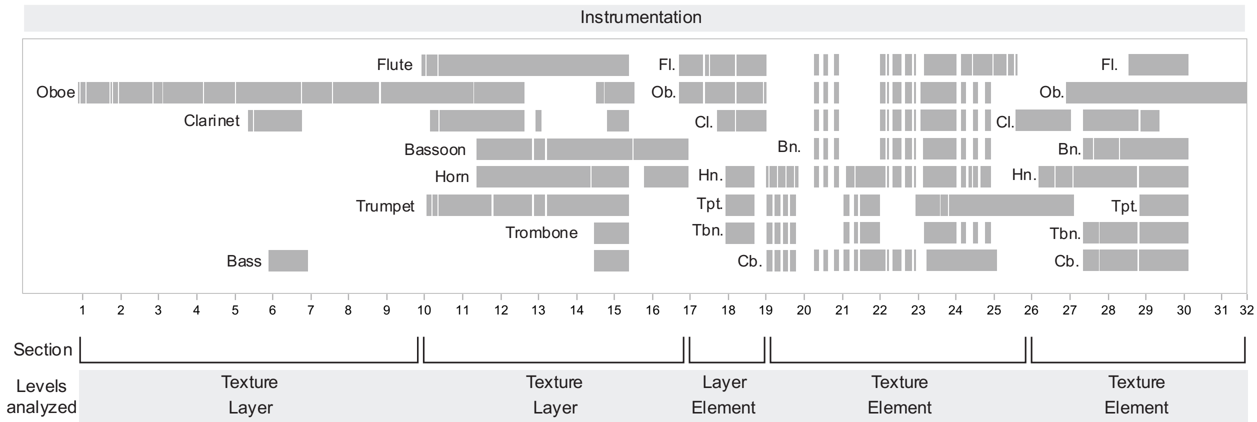


Figure 5-1. Overview of levels analyzed in each section of first movement of Octandre, aligned with instrumentation graph.

My reading of the first movement of *Octandre* integrates these three levels of analysis, which squares with Varèse's own understanding of timbre in his music. Figure 5-1 provides an overview of the chapter's approach, displaying an instrumentation graph of the movement divided into formal units, with an indication of the levels analyzed for each section. I begin with an analysis of the oboe solo (mm. 1–9), first at the level of the texture, revealing the overall brightness of the oboe in multiple recordings, and then at the level of the layer, revealing the relative brightness of the oboe, clarinet, and bass. I then proceed to a texture- and layer-level analysis of mm. 10–16, revealing a gradual darkening of the overall sound in tandem with changes in the span and distribution of concurrent layers. I integrate a layer-level analysis with an element-level analysis of the transition in mm. 17–18, showing how different analytical orientations yield different insights about timbral interactions between layers. In the next section of the piece (mm. 19–25), I quantify contrast by examining (a) shifts of the overall texture-level centroid at multiple time scales, and (b) the spacing of individual instrument centroids within the contrasting verticalities. The final section of the movement (mm. 26–32) is treated to an element-level analysis in conjunction with an overall texture-level analysis. I conclude by considering some large-scale trends in the timbral organization of the movement, pointing out differences that distinguish these sections from each other, and highlighting a basic timbral gesture ("bright-dark-bright") that appears at multiple temporal scales in the movement.

***Octandre*, I. mm. 1–9**

The piece begins, as many of Varèse's pieces do, with a monophonic line given to a solo instrument. It is worth pointing out that in solo textures such as this one, there is no meaningful distinction between the three levels of measurement: since only one instrument is sounding—in this case, the oboe—there is only one "level," and it is simultaneously an element, a layer, and the whole texture. I begin my analysis by examining the timbral fluctuation of the opening four measures of the oboe solo in five realizations, following a methodology that resembles that of the previous chapter.[6] First, I divide each audio recording into note-level snippets and compute the centroid of each one (the segmentation is shown in Figure 5-2a). Then, for each snippet, I plot the average centroid and standard deviation of its five realizations. The results are plotted in Figure 5-2b. Figure 5-2c reveals the overall average centroid of each recording, presented in order from brightest to darkest overall: while the range of brightness values in each recording is

[6] Four recordings were consulted: Robert Craft (1960, Columbia MS 6146), Riccardo Chailly (1998, Decca 460 208-2), Christopher Lyndon-Gee (2001, Naxos 8.554820), Arthur Weisberg (1972, Nonesuch H-71269). In addition, a MIDI version was realized using an orchestral sample library (Garritan Personal Orchestra 5).

comparable, the absolute brightness level varies.[7] While Figure 5-2c reveals a wide range of possible realizations, 5-2b gives us its general "fuzzy" spectral positioning—the average centroid across all recordings is shown to be 1505 Hz.[8]

Further, Figure 5-2b reveals a subtle peak in average brightness in segment 17, corresponding to the oboe's B4. This is by no means the highest pitch in the passage (in fact, it is roughly the midpoint), but nevertheless it yields the highest spectral centroid. All performances appear to place a distinct accent on this note, owing primarily to its position on a strong beat (rare in this opening solo) and its relatively long duration compared to the notes leading up to it. This accent brings about a local brightness peak in all recordings.[9]

---

[7] This variance is likely due to global differences in the sound of the recording rather than differences in the way the oboe part was performed: things like microphone placement (e.g., bringing out keyclicks), acoustics of the space, equalization, etc. will impact the overall brightness level of the audio. This fact reaffirms the importance of emphasizing brightness *relationships* over absolute centroids.

[8] As I described in Chapter 4, the purpose here is not necessarily to make analytical claims about the recordings per se, but rather to get a sense of the general brightness of the "piece itself" by taking an average of its many possible realizations.

[9] It is possible that the pitch B4 is simply a brighter sound on the oboe, even without dynamic emphasis. Support for this is found in Siedenburg and Saitis (2020), who provide a chart with centroid data of the oboe samples in the Vienna Symphonic Library; even when the performed dynamic is stable, there is a spike in brightness around the note B4/C5.

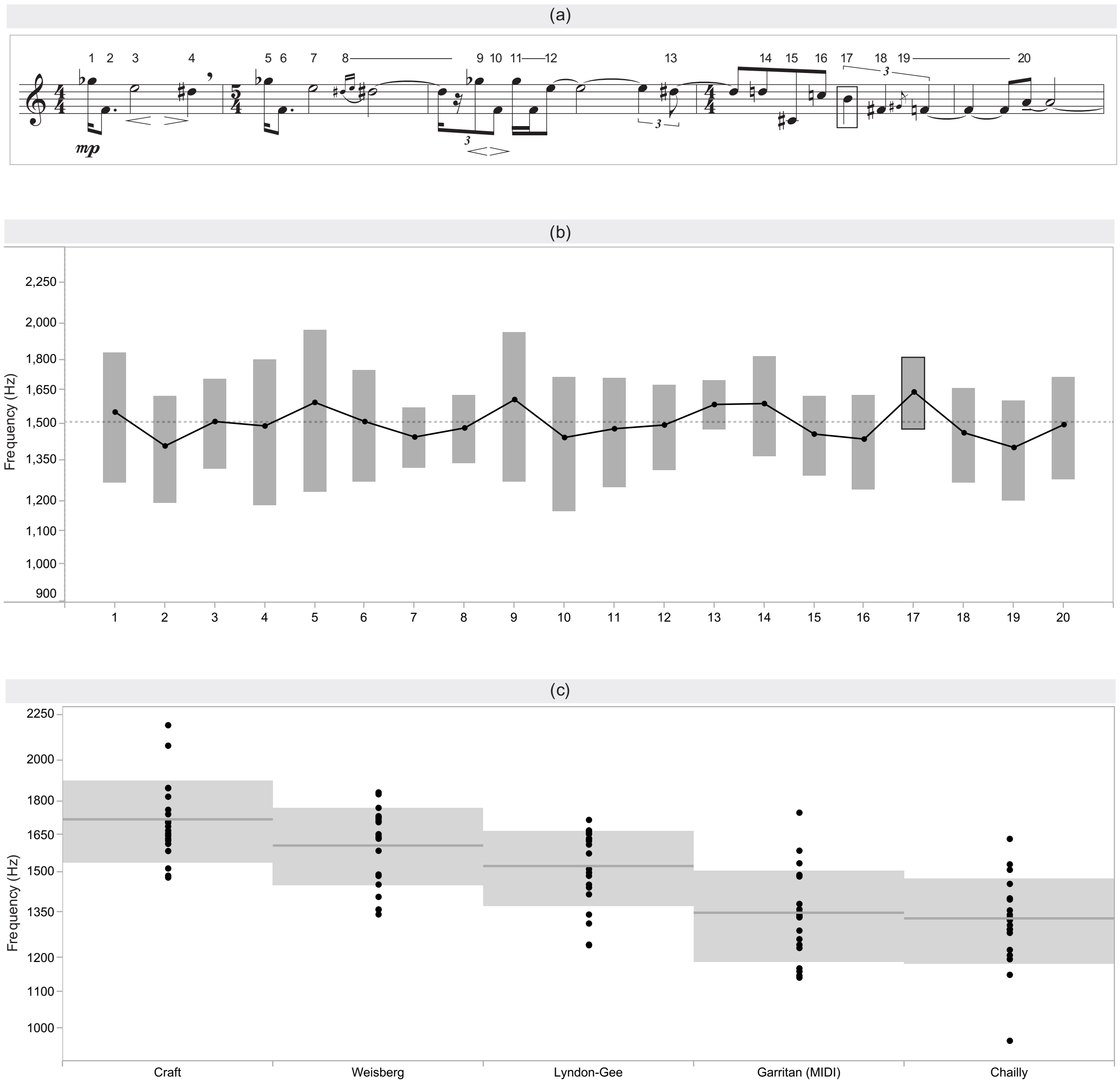


Figure 5-2. (a) segmentation of oboe solo roughly by note; (b) average brightness of each note across five recordings, with standard deviation of recordings. (c) oboe solo notes broken down by recording, with average note-level centroid and standard deviation given.

Recognizing that the oboe solo sits in a particular spectral region is useful for comparing the solo to the polyphonic music that follows, and to other instrumental solos that occur later in the piece. Let us briefly explore this latter comparison first (Figure 5-3). The piccolo solo that begins the second movement repeats continually the same G♭5 that features so prominently in the oboe solo. The piccolo's timbre, though, renders this solo brighter than the previous one (Avg. SC=2000 Hz). The third movement, on the other hand, begins with the double bass, held over from the previous movement, sounding out its B2, which is darker than the previous two openings (Avg. SC=808 Hz). This establishes a timbral process connecting the three openings: it is as if the first movement begins in the spectral middle-ground between two extremes, and each subsequent movement begins by visiting one of those extremes.

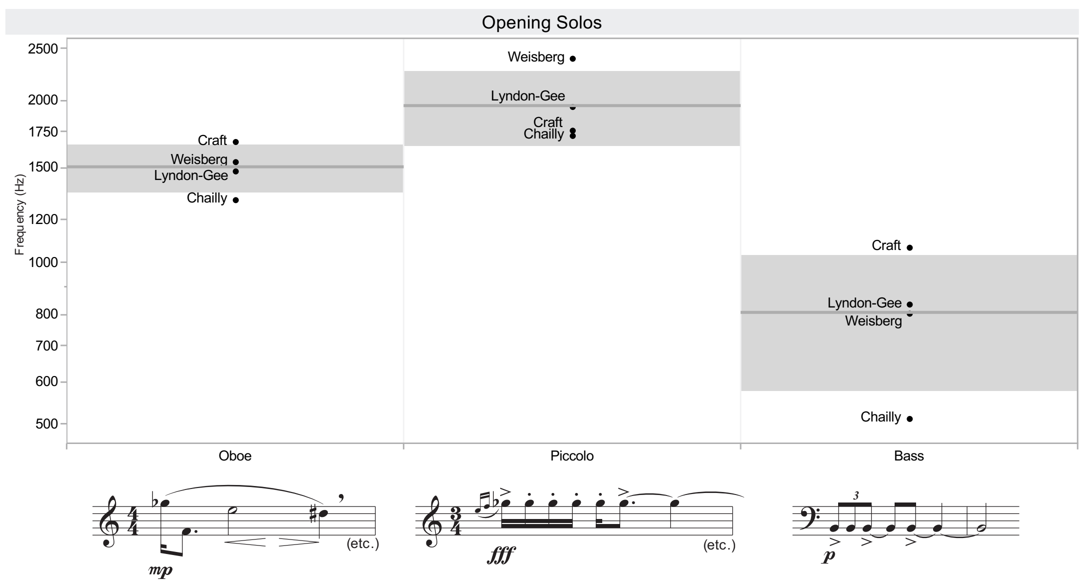


Figure 5-3. Average centroid per recording of the solos that open each movement of Octandre. The second movement (piccolo) begins brighter than the first does (oboe). The third movement begins with the darkest overall timbre of the three (bass).

Returning now to the opening of the first movement: the oboe concludes its opening sentence-like statement with a sustained A4 and is then joined by the clarinet in m. 5 and double bass in m. 6.[10] The presence of these instruments, rather unsurprisingly, lowers the overall centroid of the audio signal. In Figure 5-4a, I show this with three data points, each representing a texture-level centroid of a section, averaged across four recordings: the first section spans the oboe solo (mm. 1–5), the second spans the entry of the clarinet through to that of the bass, and the third spans the bass entry to the end of m. 6. The overall centroid between the second and third does not change drastically: when the darker bass is added to the texture, the oboe ascends into a brighter range to compensate.

This texture-level approach (5-4a) allows us to consider the broad acoustic outcome of the clarinet-bass addition; an alternative reading, however, examines the spectral layout of the three instruments (oboe, clarinet, and bass) as concurrent layers. As we can see in Figure 5-4b, the oboe remains at the "top" of the texture (SC=1587 Hz), with the clarinet in the middle (695 Hz) and the bass at the "bottom" (480 Hz). The overall *range* of the layers (i.e., the distance between oboe and bass) spans a large centroid ratio (SCQ=3.31), and the three layers are unclustered, yielding a maximally "transparent" texture (i.e., with a layer-level agglomeration score of 0), albeit with the clarinet and the bass slightly closer together.

[10] Babbitt (1966) observes that the clarinet motive here (C♯ B B♭) functions somewhat like an "answer at the fifth" to the opening motive stated by the higher notes of the oboe (F♯ E E♭). Indeed, this three-note motive and the chromatic tetrachord it is extracted from completely saturate this movement.

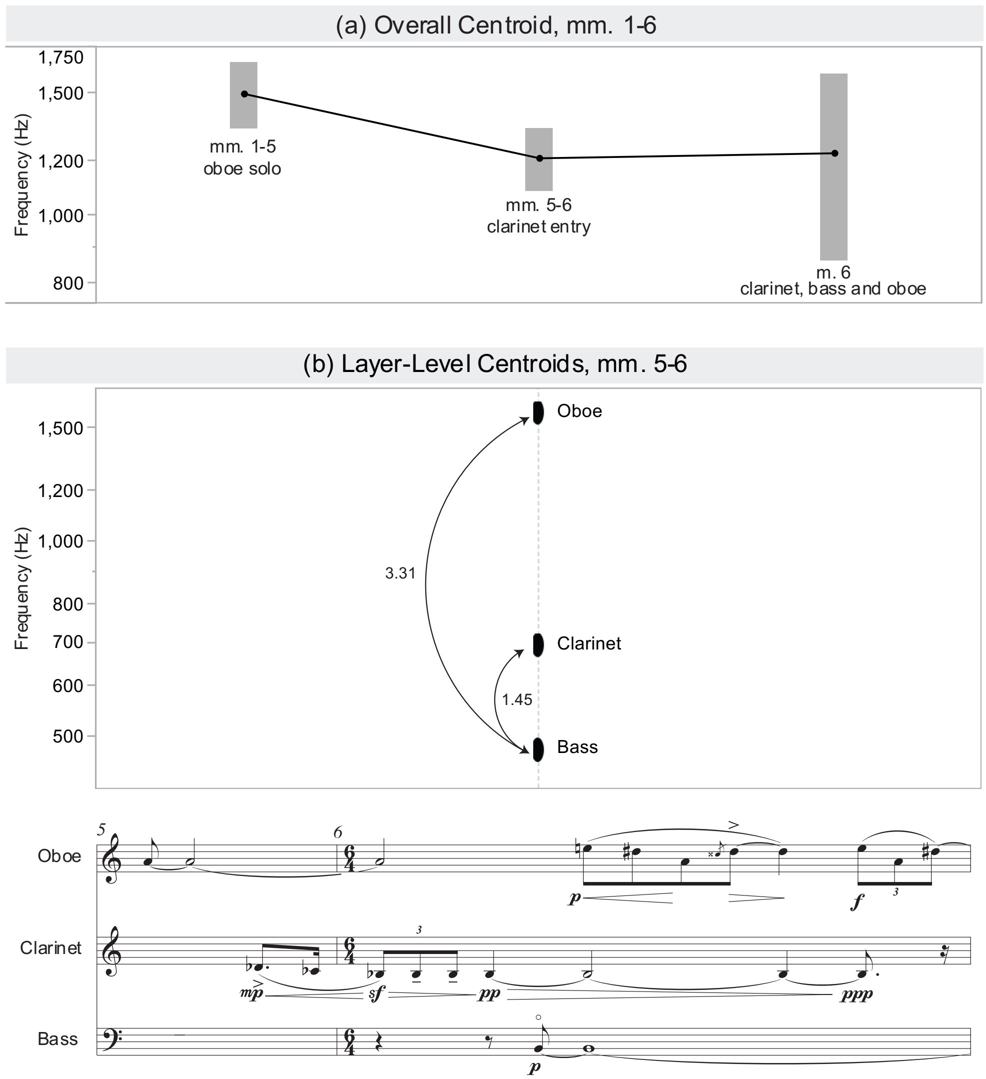


Figure 5-4. (a) The texture-level centroid darkens at the entry of the clarinet and stays at roughly the same level when the bass joins. (b) Layer-level analysis of mm. 5–6, with SCQs labeled, shows “non-blending” of the three instruments, with slight clustering of clarinet and bass.

Octandre, I. mm. 10–16

The bass and clarinet each recede to silence in m. 6, and the oboe continues alone, eventually shooting upward to a dramatic peak of G6 in m. 10 and thereby signaling the onset of a new formal section. I divide the ensuing section (mm. 10–16) into five segments on the basis of pitch and textural changes. A score reduction with my segmentation is given in Figure 5-5. The five segments are labeled with lower case Roman numerals along the bottom of the score. Colors on the score represent a rhythm-based partitional reading of each segment, described in more detail below.

As the oboe's arrival pitch (G6) is held, a trio of flute, clarinet, and trumpet joins in, each instrument loosely imitating the oboe's upward trajectory to a different sustained pitch, forming a piercing four-note chord in m. 11 (A♭5–A5–G6–C♯7). A new segment (ii) begins on the last beat of m. 11, when three of the four instruments (all but trumpet) move to new pitches: the flute leaps down a fourth (G♯6) and the clarinet and oboe move together to a dyad (F5–G♭5), marking the first instance of onset synchrony in the piece. A duo of bassoon and horn enters promptly—again with onsets synchronized—on a lower sustained dyad (C4–B4). With the trumpet still holding its arrival pitch from segment i (A5), the result is a six-note sonority that reduces in intensity and then freezes momentarily under the fermata in m. 12. The next segment (iii) begins when activity resumes: the bassoon disconnects from the horn and moves through a series of dyads in sync with the trumpet, leaving the horn to sound out its C4 in a separate layer. The texture here is thinner: the oboe is absent altogether and, after a brief interjection in m. 13, the clarinet withdraws too. The fourth segment (iv), by contrast, involves all eight instruments, each appearing in rapid succession at a fortissimo dynamic, resulting in a massive sonority that grows toward a powerful arrival in m. 15. The section terminates in segment v with the bassoon plunging to its low register, prompting the horn to rejoin with a similar downward dive.

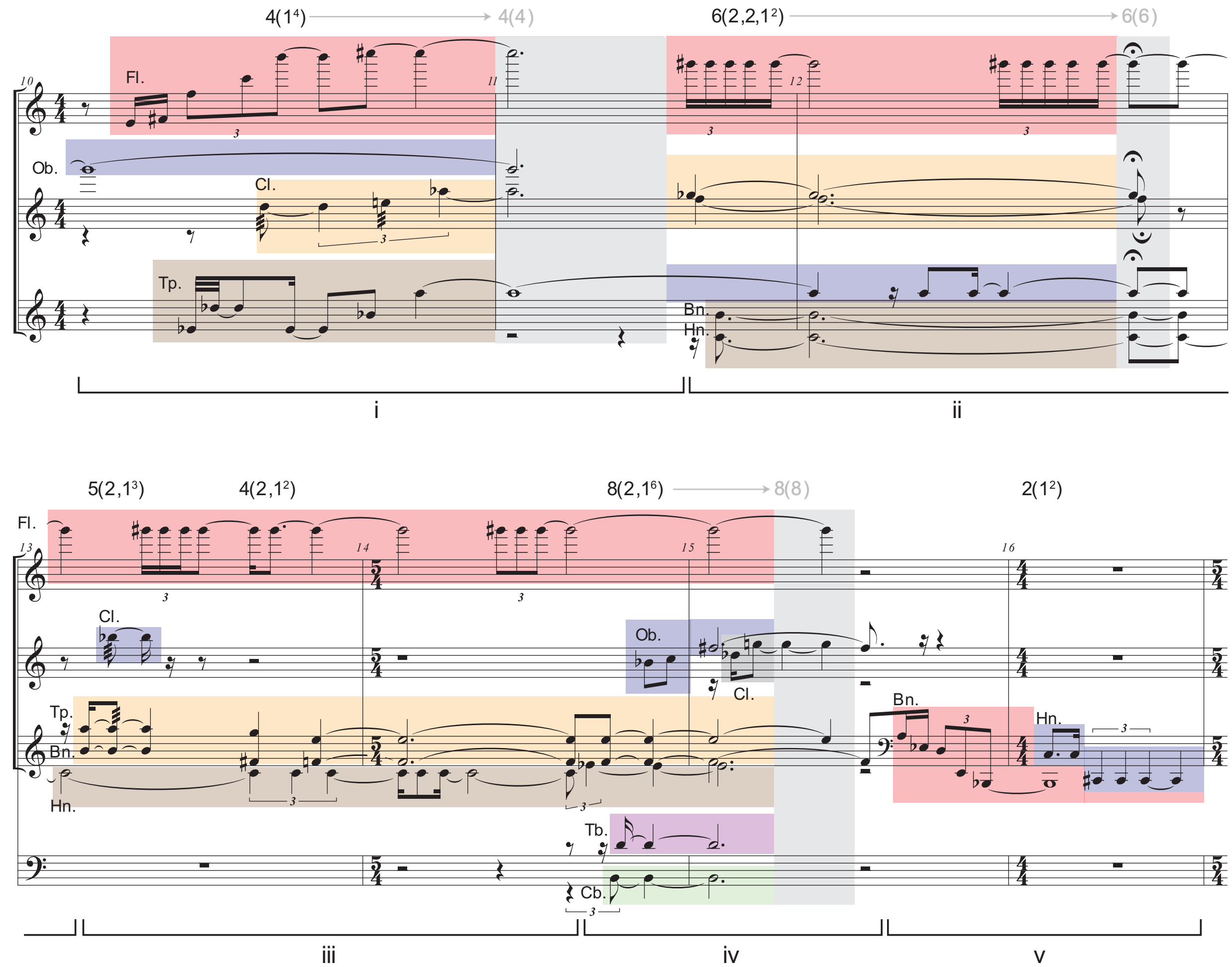


Figure 5-5. (a) Reduced score of mm. 10–16 of *Octandre*, I., with formal segmentation and partitional analysis. The passage divides into five segments. Uncoordinated elements cohere into unified layers over time, despite a lack of onset synchrony.

Colors in Figure 5-5 show my partitional reading of this passage—delineating the music into distinct textural layers—which is mirrored by partitional notation above the score. By and large this partitional analysis relies on onset synchrony to group elements together and distinguish layers from each other. However, the grey rectangles (mirrored by grey partitional expressions) represent a unique textural situation I call *layer convergence,* which occurs three times in the passage (and elsewhere in the piece): elements whose onsets are *not* coordinated rhythmically nevertheless appear to "fuse" into a single textural layer over time. By m. 13, for example, four sustained elements have converged into a unified layer, despite each beginning its life at a different point in m. 12 or earlier.

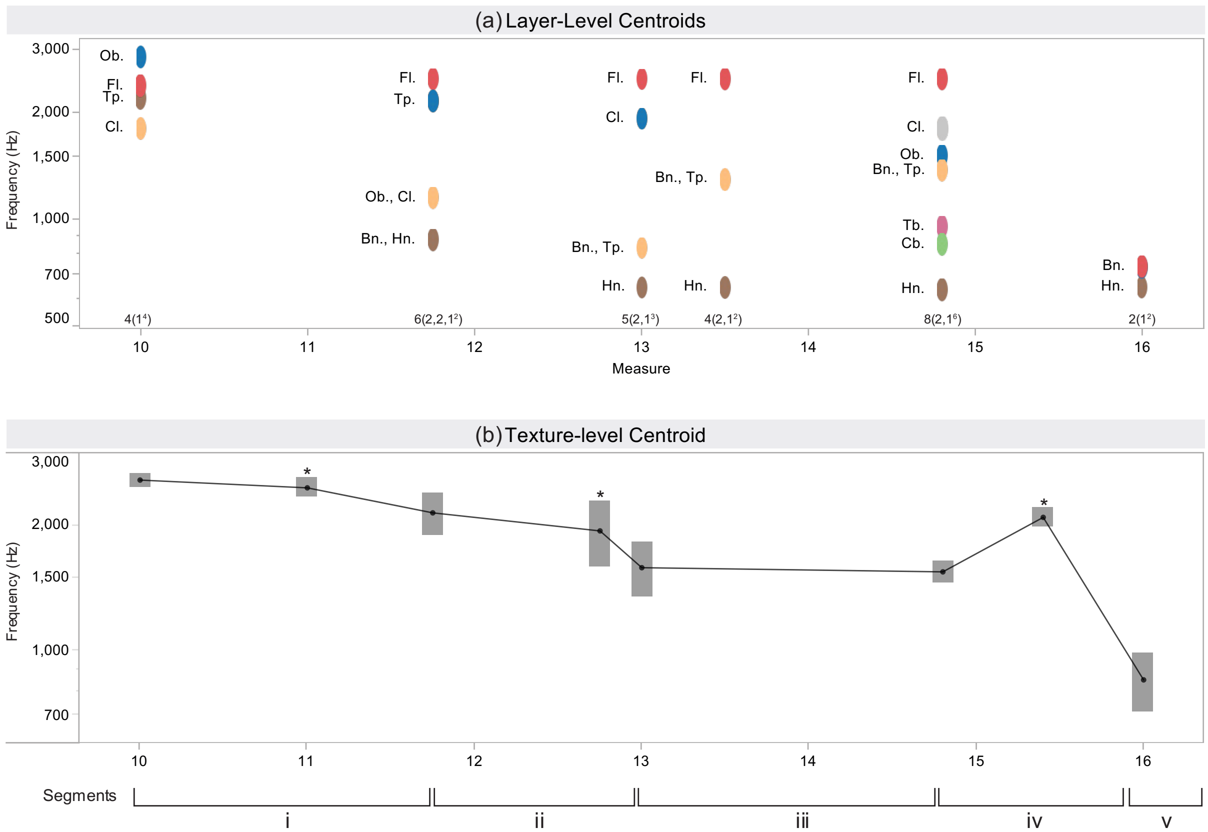


Figure 5-6. (a) Layer-level centroids of mm. 10–16. Colors match points to layers in Figure 5-5. (b) Texture-level centroids of same passage. Asterisks indicate moments of fusion that occur in the absence of onset synchrony (corresponding to the grey sections in Figure 5-5).

A layer-level analysis, showing the spectral layout of these instrumental layers is given in Figure 5-6a. Each layer revealed by the onset-based partitional analysis (in Figure 5-5) is plotted as a point in 5-6a, with vertical distances representing perceptual intervals of brightness.[11] The layer-level analysis reveals a gradual process: over time the SCQ between

[11] As described in Chapter 1, my analysis of isolated layers and elements relies on MIDI data performed by a sample library (Garritan Personal Orchestra 5). To get each layer

the lowest and highest layers increases and new layers fill in the space between them, enacting a saturation of spectral space in segment iv, before a darker, low-span texture appears in segment v. There is a clear symmetry in the passage: four widely spaced textures in the middle (mm. 11–15) are flanked by two more spectrally compact textures (in mm. 10 and 16).

This layer-level reading (5-6a) observes strictly the onset-synchrony of the passage; multiple instruments only get collapsed into a single point on the graph if they begin at precisely the same time. Thus, it does not reflect the layer convergence described above, wherein rhythmically distinct layers cohere over time. This is captured instead by the texture-level reading in 5-6b, which shows the overall centroid of the entire texture.[12] Asterisks mark the points corresponding to these moments of fusion in the absence of onset synchrony. From this vantage point, we can also see that the five segments of the passage enact a gradual darkening procedure, moving toward the bassoon's low-register descent.

This passage (mm. 10–16) features the first instances of onset synchrony: while the opening measures—the oboe solo with added clarinet and bass—exhibits a total absence

---

centroid, I took the average of four instrumental realizations of the MIDI (e.g., for the clarinet layers, I analyzed the "Clarinet Player" 1, 2, and 3 samples, along with the "Clarinet Solo" samples).

[12] Since a layer-level analysis of a single unified layer would yield a single centroid, it is effectively the same as a texture-level analysis.

of rhythmic alignment, we now find pairs of instruments (duos) beginning together. The trend continues in the passage immediately following, which serves as a transitional section. We find in m. 17 a synchronized duo of flute and oboe, and then—in m. 18—a case of three onset-synchronized instruments (flute, clarinet, and oboe) pitted against another synchronized trio (horn, trumpet, and trombone).[13] More instruments are beginning to coordinate with each other, and textural layers are therefore growing denser. This trend of increasing coordination and density continues further, leading to an extended passage characterized by a substantial amount of onset synchrony, now with the densest layers yet: beginning in measure 19 ("Lourd et sauvage"), a synchronized trio (trumpet, trombone, and bass) and solo horn alternate with a synchronized sextet of flute, oboe, clarinet, bassoon, horn, bass (m. 20). This large-scale densifying process leading to the arrival of the sextet in m. 20 is summarized by the gradually increasing agglomeration scores shown in Figure 5-7. This indicates how many *pairs of elements* are rhythmically coordinated in the texture (described in more detail in Chapter 3).

[13] More precisely, the wind trio is loosely synchronized. The clarinet joins aligned with the flute, in m. 18, but breaks off slightly from the other two. Synchronization completely dissolves on the last beat of m. 18. This is thus what Wallace Berry calls a "heterorhythmic" combination (i.e., the strands are *somewhat* aligned rhythmically).

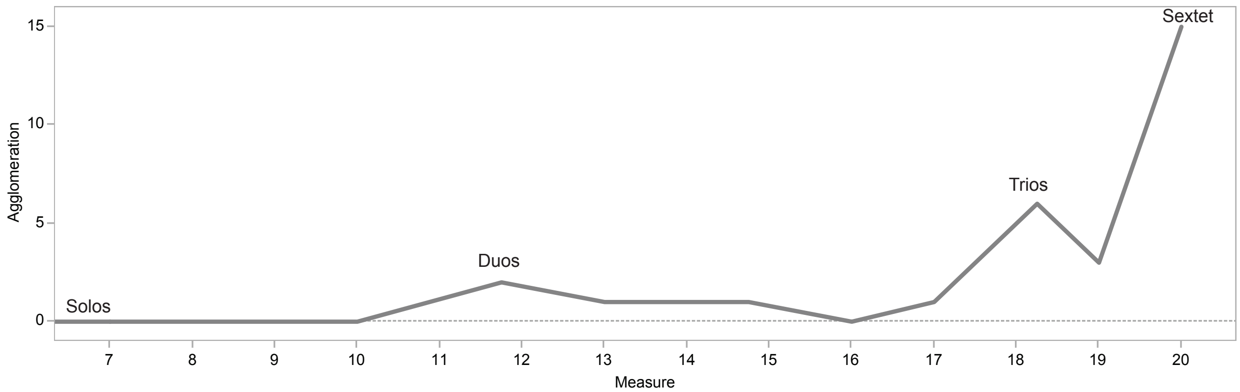


Figure 5-7. Increasing agglomeration trend for Octandre, I., up to m. 20: textural layers are getting larger. Annotations indicate the largest textural layers present at each point.

The transitional passage from mm. 17–18 is plotted at two levels in Figure 5-8. In 5-8a, a layer-level analysis shows an expansion outward in spectral space: the flute-oboe duo brightens slightly with the addition of the clarinet in m. 18 (SCQ=1.08), and the simultaneously added brass trio sits darker than the flute-oboe duo by precisely the same amount (SCQ=1.08). Figure 5-8b shows the element-level make-up of these textures. The proximity of the two elements in the flute-oboe duo (m. 17) suggests a nearly homogeneous combination. Further, the two layers in m. 18—wind trio vs. brass trio—form an overlapping configuration not apparent in the layer-level analysis: the brass trio, while darker overall, has a much wider span (SCQ=2.81) than the wind trio (SCQ=1.27). And the trombone centroid in the brass layer overlaps with the clarinet and flute centroids in the wind layer, suggesting timbral interaction between elements of different layers. Here, different analytical orientations yield different insights: if we choose to see layers as coherent wholes, then the wind trio is altogether brighter than the brass trio, but if we instead view layers as collections of elements, the wind trio is *nested within* the brass trio, with certain elements overlapping.

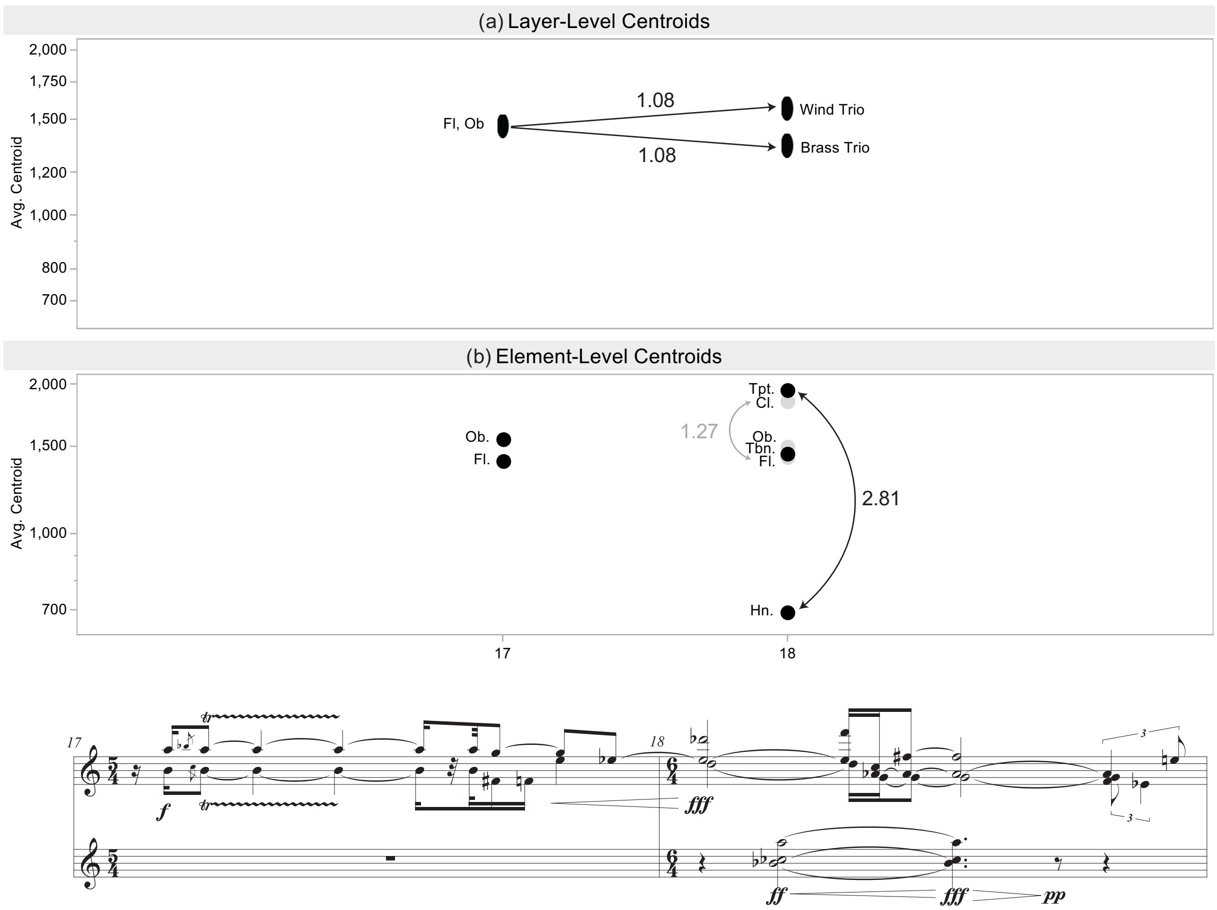


Figure 5-8. (a) Layer-level analysis showing outward expansion from a duo layer to two trio layers. (b) Element-level analysis showing narrow wind trio nested within dispersed brass trio.

This pair of synchronized trios, pitted against one another in m. 18, appear to exemplify Varèse's notion of simultaneous sound masses. Such is the opinion, for example, of Erickson (1975) and Smoot (1986).[14] Interestingly, both these authors implicitly suggest that Varèse's "sound masses" necessarily involve synchronous onsets, and at least three

[14] Erickson (1975) refers to the lower brass trio as a "fused ensemble timbre"—his chosen synonym for "sound mass." Smoot (1986) gives the same chord in a list of sound masses in *Octandre*. Each mass involves at least three instruments and exhibits synchronized onsets.

instruments. On this view, this rather fleeting transitional passage (mm. 17–18) in fact marks a major juncture in the piece, as it is the very first instance of a “sound mass”: the earlier verticalities arrived at via layer convergence in mm. 10–16 are not sound masses, on this view, and neither are the synchronized duos found within them. This strikes me as an unnecessarily restrictive definition of a concept that Varèse treats as ubiquitous in his music; a more inclusive definition, more in line with Varèse’s usage, would encompass the synchronized duos of mm. 10–16 and the moments of layer convergence examined previously. This broader definition thus treats the passage from mm. 10–16 as involving simultaneous masses (duos) and solos growing together to form larger masses, which in turn separate again into smaller entities, a process that mirrors Varèse’s image of “penetration” and “repulsion” of sound masses in space. This view also allows the simultaneous trios at m. 18 to be understood as one part of an accumulative process of agglomeration described above, wherein the sound masses of the piece are gradually becoming more synchronized, with more and more members being grouped together into layers, leading to the heavily synchronized sound masses of the “Lourd et sauvage” passage that begins in m. 19.[15]

---

[15] Lalitte (2011) points out that Varèse’s first encounter with the term “sound mass” (*mass sonore*) was in Helmholtz’s *On the Sensations of Tone* (1863/1875), wherein it refers to the overtone spectrum of isolated instrumental tones. So, while I won’t pursue this line of thought, it seems plausible that Varèse considered even *solos* to be a kind of “sound mass.” This finds support in Reynolds (2013). Anderson (1984) refers to solo layers as “planes,” while duos and beyond are “masses.”

Octandre, I. mm. 19–25

Let us now examine the "Lourd et sauvage" passage at two levels of detail (texture and element). This section is an example of what Wuorinen calls the "juxtaposition of differentiated elements" in Varèse's music (Julius, 1978): measure-long segments of dissimilar character alternate and propel the music forward through their antiphonal contrast.[16] We will attempt here first to grasp the degree of differentiation between the opposed segments via texture-level centroids at multiple temporal scales, before applying an element-level analysis.

Figure 5-9a gives a score reduction of the first four measures of this section (mm. 19–22). I break the passage into four segments (by measure) labeled with lower-case Roman numerals, and several subsegments labeled with lower-case letters. Segment i alternates between two sounds: a brassy low-register chord (labeled a), and a repeated C♯4 in the horn (labeled b). Segment ii counters with a woodwind-heavy chord—the synchronized sextet at the end of the agglomeration line in Figure 5-7—which repeats itself three times. The material of segment i then returns in segment iii, this time with the horn slightly less coordinated with the low-register chord. This segment includes a new subsegment, labeled c, wherein the horn and low-register chord execute a massive crescendo. Segment iv elaborates on the woodwind-heavy segment ii, alternating between two different six-note chords.

[16] Anderson's discussion of the first four measures of this passage refers to the contrasting segments as "sound masses separated in time" (quoted in Moura, 2004).

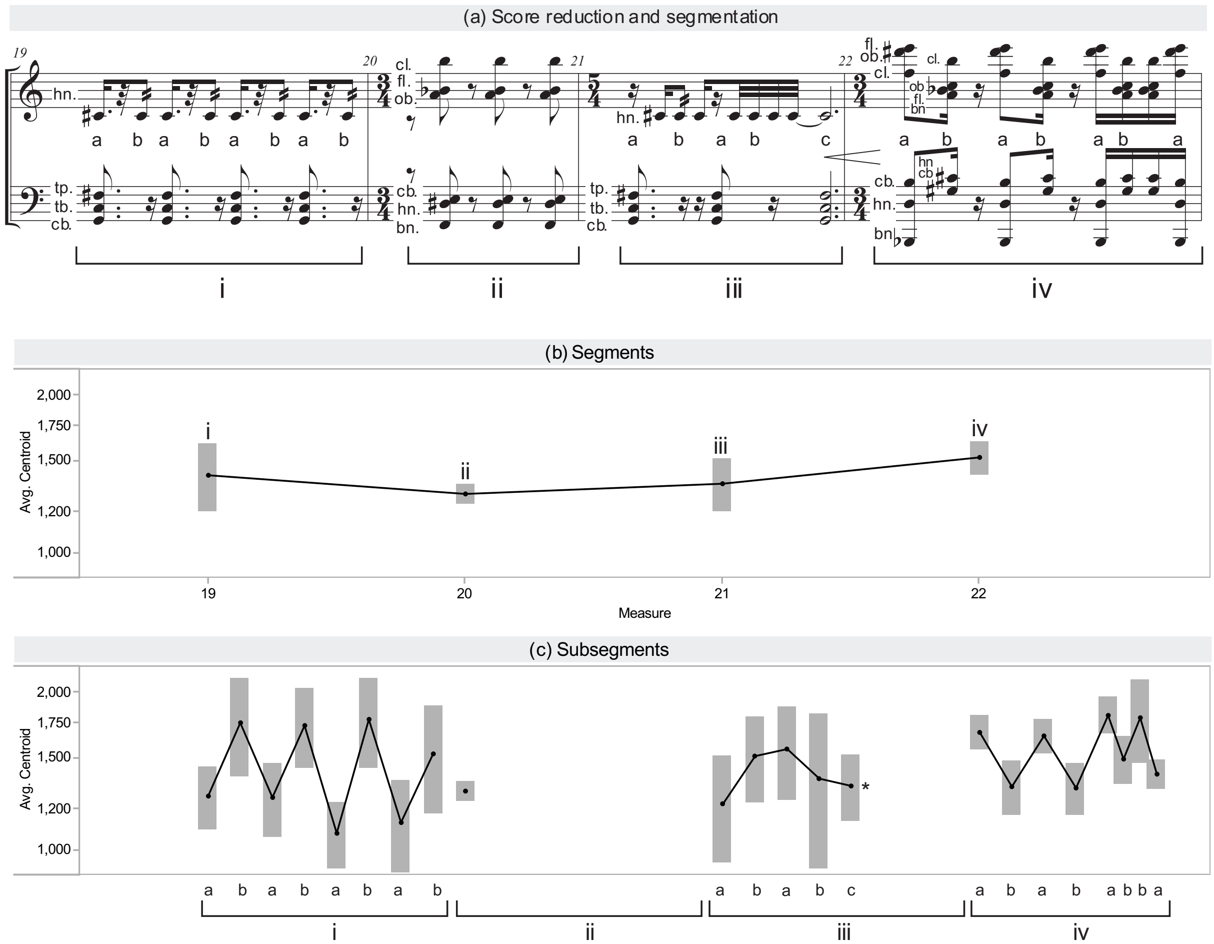


Figure 5-9. (a) Score reduction and segmentation of mm. 19–22. (b) Texture-level analysis (overall centroid) of segments in mm. 19–22 of *Octandre*; (c) Texture-level analysis (overall centroid) of subsegments of same passage. Grey bars indicate standard deviation of the four consulted recordings.

A texture-level analysis of this four-measure passage is shown in 5-9 (b and c). At the level of the segment (5-9b), interestingly, the overall shift in brightness between adjacent segments is not particularly high: while segments i and ii, for example, each have a very different timbral character, their overall spectra nevertheless have similar centers of gravity. Thus, the perceived timbral difference (i.e., between i and ii) is likely better understood by examining the constituent *elements*, an approach that will be undertaken below. The subsegment analysis (5-9c) reveals internal contrasts within segments—in segment i we find alternation between relatively dark (a) and bright (b) subsegments. This situation is reversed in segment iv, wherein the first subsegment (a) is brighter than subsegment b.

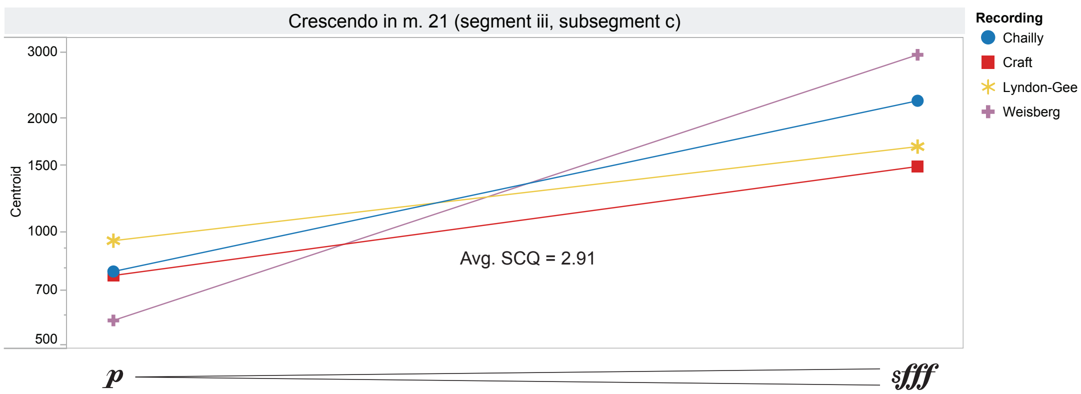


Figure 5-10. Increased playing intensity in m. 21 correlates with increased texture-level brightness (Avg. SCQ = 2.91).

The foregoing analysis fails to capture an important surface-level timbral change that occurs as a result of the dynamic swell in segment iii, subsegment c (labeled with an asterisk in 5-9). As Babbitt (1966) explains, Varèse's crescendo markings are first and foremost a means of timbral transformation: "crescendi ... produce not what can be most accurately described as a change in loudness of a fixed sonority, but a continuous alteration of the number, relations, and densities of the partials of the total spectrum." Thus, we must zoom in further to examine surface-level contrast within individual notes or verticalities. Figure 5-10 shows the results of such an analysis, broken down by recording. We find large spikes in brightness in all recordings, albeit to different degrees: the most dramatic crescendo here is in Weisberg (SCQ=5.07), while the least dramatic is Lyndon-Gee (SCQ=1.77).

Texture-level centroids capture well the contrast between adjacent subsegments in mm. 19–22, but, as mentioned previously, fail to reflect the perceptual difference between adjacent segments (e.g., i and ii). An analysis of concurrent elements, however, will reveal further detail about the internal timbral structure of these segments: differences in the spread and distribution of element-level centroids will give us a sense of the overall color of each verticality and a useful means by which to compare it to other verticalities.

Figure 5-11 shows the element-level centroids comprising segments i, ii, and iv (mm. 19–22) and their subsegments.[17] The shift in brightness between segment i-a and i-b is further clarified at this level: the crescendo, given to the trumpet, trombone, and double bass in subsegment a, brings each instrument's centroid notably higher for subsegment b. In addition to sounding brighter overall (as seen previously in Figure 5-9), the effect is of a contraction of the overall *span* of the trio of instruments that accompany the horn layer (shown with grey points in Figure 5-11)—the range of the lowest and highest element-level centroids in the trio becomes narrower (SCQ=1.55). In both subsegments, the trumpet is the brightest timbre and is separated from the rest of the ensemble due to the brassy nasality of its extreme low register.

---

[17] As with the layer analysis in Figure 5-5, the data points here represent an average of four realizations of each MIDI note. The elements in segments i and iii are for the most part identical, so I have chosen to omit segment iii from this visualization.

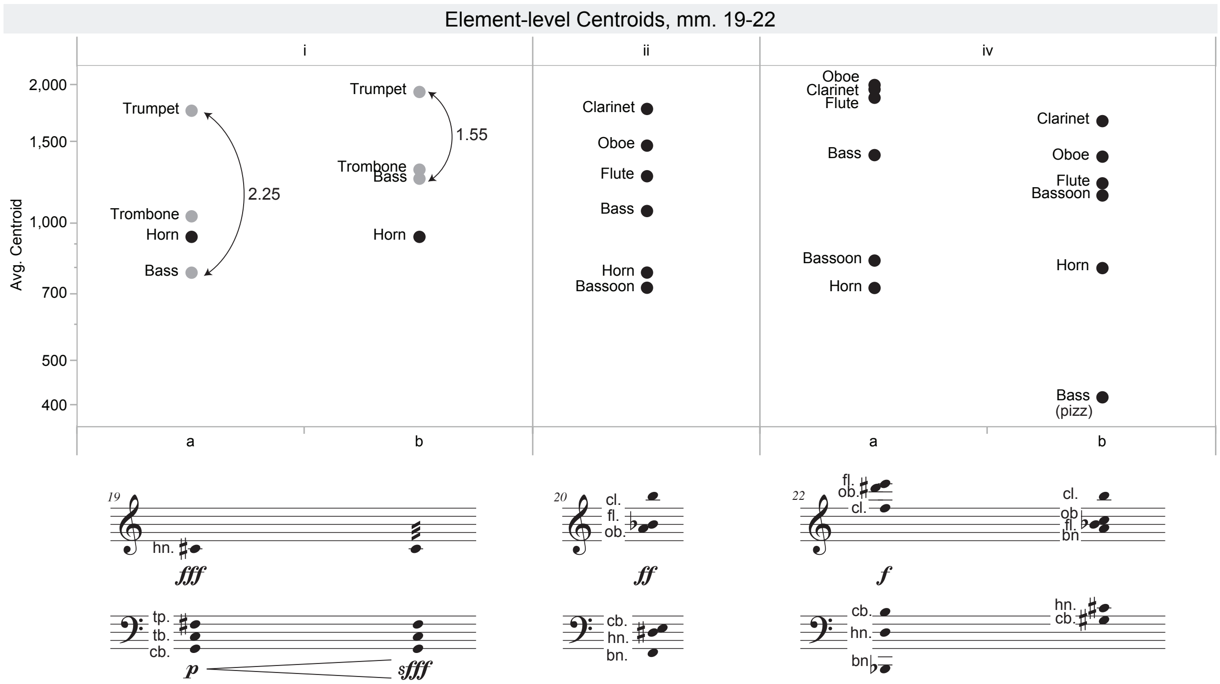


Figure 5-11. Element-level analysis of mm. 19–22. The span of centroids becomes narrower in segment i subsegment b. In segment ii, the distribution of centroids is much more even than in the other segments.

The contrast between segments i and ii can be explained in terms of centroid distribution: segment i is visibly more *uneven* than segment ii. In Chapter 2 we defined *unevenness* as the standard deviation of adjacent SCQs in a sonority—an unevenness of 0.00 indicates that all the SCQs between adjacent elements are the exact same size. In this case, the unevenness score for segment i (subsegment b) is 0.23. Segment ii, however, is notably lower (0.10). On a similar note, segment i-b groups into three distinct clusters (trumpet, trombone + bass, horn), while segment ii appears to exhibit no obvious clustering. Interestingly, the spacing of the *pitches* in these segments works in precisely the opposite way: while segment i is rather uniformly spaced (in stacked fourths and fifths), segment ii exhibits a more clustered distribution of pitches.

In segment iv-a (m. 22), the three woodwind centroids at the top of the sonority—flute, oboe, and clarinet—are positioned very close to one another, virtually overlapping. This suggests a perceptual overlapping of these elements, resulting in a unified, nearly fused trio that stands distinct from the other three elements in the sonority (bassoon, horn, and bass). This approach builds on the "fused ensemble timbre" concept that Erickson (1975) defines with reference to Varèse: in this instance (iv-a), onset synchrony indeed induces an overall sense of perceptual integration, but we can further analyze the structure of the mass and show which constituent elements appear more fused than others and why.

The next three measures (mm. 23–25) play a transitional role in the textural structure of the movement. The measure-to-measure antiphonal pattern continues: in m. 23, the trumpet gives its version of the opening oboe (and clarinet) motive, accompanied by another case of *layer convergence*, wherein eight elements enter separately (in four synchronized pairs), but then perform a massive crescendo together. In the next measure (m. 24) the woodwind-heavy material from segments ii and iv returns to accompany the trumpet, holding its note over from the previous measure. Then, having reached a peak of synchronization (with the three instances of the sextet in mm. 20, 22, and 24), the music suddenly breaks apart into several distinct solos in m. 25 and 26. In the context of the whole movement, then, the passage from mm. 19–25 represents a kind of textural crux, carrying out the unification of individual elements into a cohesive mass, which then dissolves again into separate layers. This dissolution ultimately continues, leading to the terminal oboe solo in m. 30. The agglomeration plot given previously in Figure 5-7 is thus roughly reversed for the second half of the piece.

In terms of onset synchrony (or agglomeration), then, the first movement of *Octandre* proceeds through a three-part structure: (1) minimal synchrony (mm. 1–18), (2) substantial synchrony (mm. 19–25), and (3) minimal synchrony (mm. 26–32). This pattern is mirrored at the level of the entire three-movement piece: the first movement has generally low onset synchrony, except for the central section analyzed above. The second movement appears to exhibit substantially more synchronization, while the third movement, with its fugato structure, retreats to a state of high dispersion.[18]

Octandre, I. mm. 26–32

In mm. 26–27, the horn, clarinet, and trumpet are all distinctly audible but not evenly spaced in terms of brightness. Despite their rhythmic asynchrony, the trumpet (F#5) and clarinet (C6) cluster together (SC = ~1855 Hz), creating a blended pair that floats above the darker horn (SC = 800 Hz), which prominently echoes the trumpet motive from three measures earlier. The result is thus two perceptual strata delineated by timbre, marking another instance of the "non-blending" concept so integral to Varèse's thinking (Figure 5-12a).

[18] Several local instances of layer convergence, such as those in mm. 10–14, can also be viewed in terms of this ternary "apart–together–apart" structure.

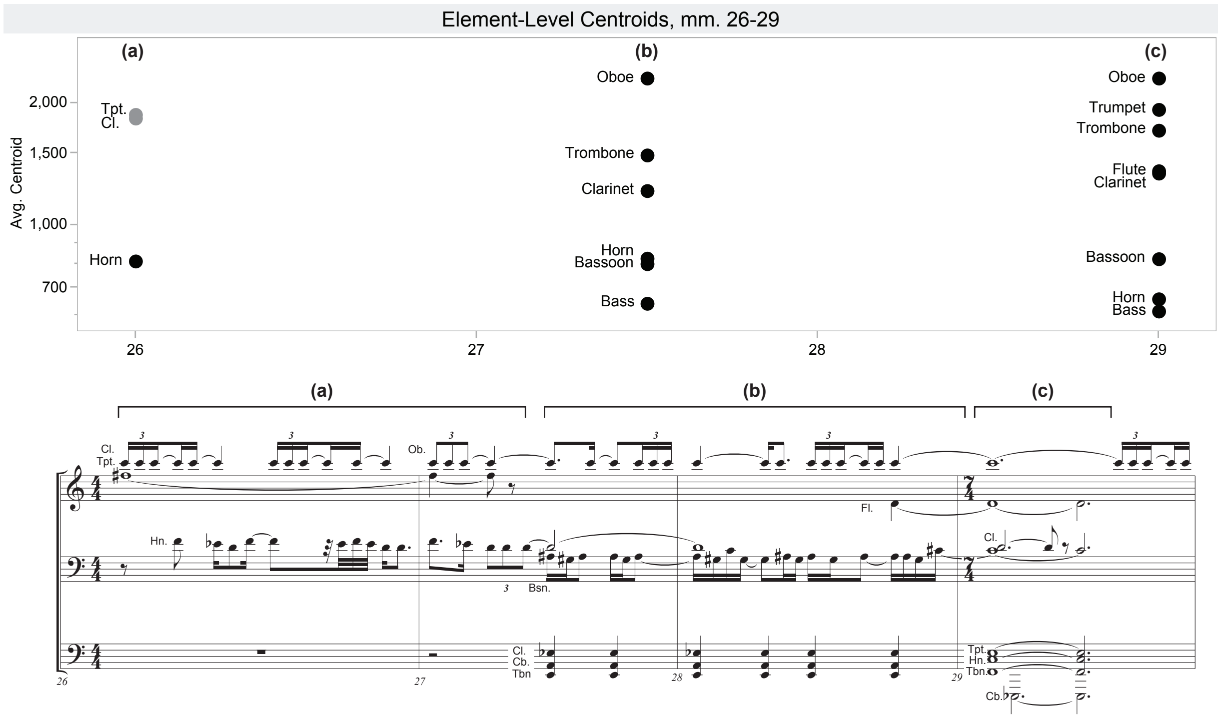


Figure 5-12. Element-level analysis of mm. 26–29. In (a), the trumpet and clarinet blend despite onset asynchrony.

This passage features an effect that is common in Varèse's oeuvre: a single note being passed around to multiple instruments. The note C6 first appears in measure 25 at the end of the flute flourish. It is immediately picked up by the clarinet and passed to the oboe in measure 27, before becoming the starting note of a final few statements of the opening four-note motive, now transposed up a tritone. The clarinet and oboe have very similar timbres at this pitch and dynamic, but the oboe is decidedly brighter. The switch to a brighter timbre appears to serve the function of bringing about more transparency in the texture that begins in the second half of m. 27 (Figure 5-12b): the oboe is bright enough that it does not overlap spectrally with any other simultaneous musical elements—indeed, it stands out from the pack, remaining clearly audible in the polyphony.[19]

The massive chord in m. 29 (Figure 5-12c) is an instance of layer convergence that suggests multiple textural interpretations. Two chord members begin precisely together on the last beat of m. 28, when the flute joins the oboe's C6 an octave plus minor seventh lower (D4). This might be viewed as a single layer involving two elements; however, given that the oboe's C6 has been heard as its own solo layer for seven beats, its momentary alignment with the flute is not powerful enough to cancel out the perception of two distinct layers. That their centroids are rather separated in spectral space does nothing to aid fusion. Also of interest is the clarinet, entering on the downbeat of m. 29 with the horn, trumpet, and trombone. We might view this as a four-part layer, but the clarinet is given a much shorter note than the rest of the ensemble (who fade out together): while its onset is synchronized with the ensemble, its endpoint is not. This is because the clarinet, playing D4, acts as a momentary unison coloration of the flute's D4; this rare unison is intended to ensure that this pitch, otherwise carried by just the flute's weak low register, will remain audible amidst the fortissimo brass timbres in m. 29. The clarinet leaves early so as not to unduly emphasize that same pitch during the dramatic decrescendo that follows. That the flute

[19] The clarinet, on the other hand, would intrude somewhat on the spectral region of the trombone, making both less audible in the process.

and clarinet belong together is supported by Figure 5-12c, wherein their centroids are shown perfectly overlapping.

Figure 5-13 shows my texture-level analysis of these three segments from mm. 26–29 (a, b, and c), along with the terminal oboe solo (d). The overall centroid during the horn motive (a) (1689 Hz) darkens at the entry of the bassoon melody with brass accompaniment (b) (1295 Hz) and remains in place for the massive chord (c) (1250 Hz), before brightening again for the oboe solo (1750 Hz). Thus, we find at the conclusion of the movement an overall "darkening and re-brightening" gesture heard earlier, both in mm. 19–22 (shown previously in Figure 5-9) and in the opening entry of the clarinet and bass (shown previously in Figure 5-3).

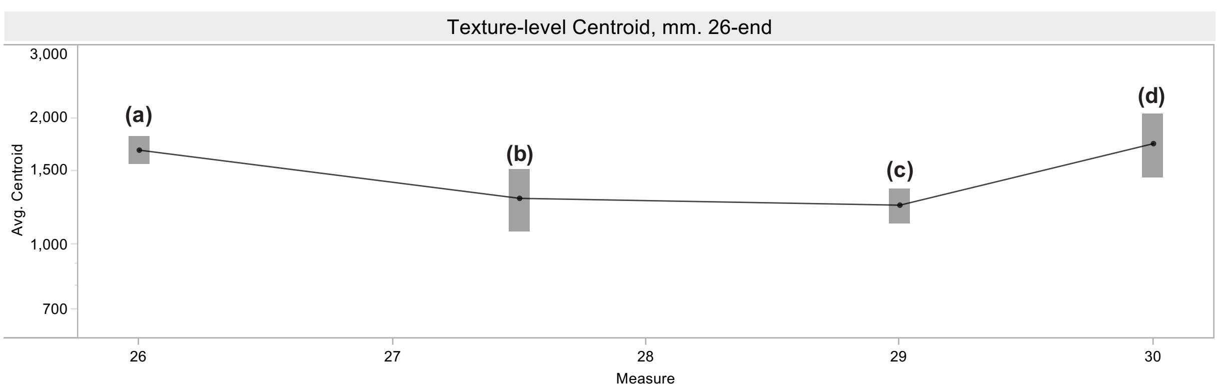


Figure 5-13. Texture-level analysis, mm. 26 to the end of the piece. Darkening and re-brightening to conclude the movement.

Timbre as an Agent of Delineation

Having explored multilevel brightness in each of the main formal chunks of the work, I wish now to conclude with a consideration of the form as a whole, with reference again to Varèse's notion that timbre is an agent of delineation—an "essential part" of form. Figure 5-

14 integrates all foregoing observations about the piece into a single visualization. The top panel (5-14a) shows the overall texture-level centroid, coordinated with an instrumentation graph below (5-14b). The bottom panels show layer-level analyses (5-14c) and element-level analyses (5-14d) of select passages (corresponding to spans of music marked in the top panel, 5-14a.

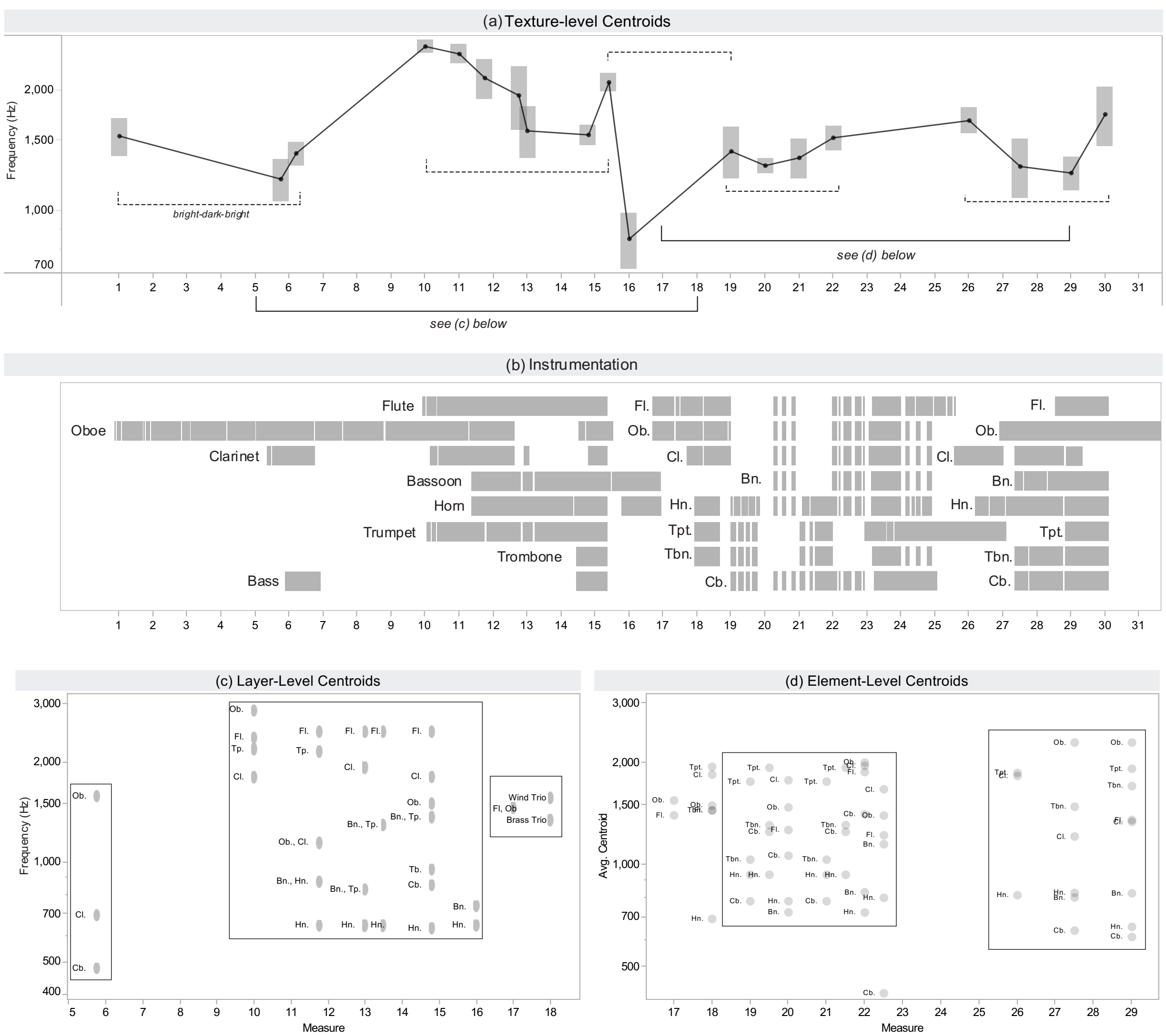


Figure 5-14. Multilevel analysis of Octandre, I. (a) Texture-level centroids of entire movement; (b) Instrumentation of the piece; (c) Layer-level analysis of mm. 5–18; (d) Element-level analysis of mm. 17–29.

I will begin by exploring how timbre "delineates" the main formal units of the piece at each level. The units in question are those covered in isolation above: (1) the oboe solo from mm. 1–9, (2) the converging chords in mm. 10–16, (3) the transition in mm. 17–18, (4) the antiphonal contrasts in mm. 19–25, and (5) the final section leading to the terminal oboe solo in mm. 26–32.[20] My goal here is not to claim that timbre is the only, or even the most important factor in the delineation of these sections; the music is shaped by a confluence of musical parameters working together. Rather I wish to determine the sense in which, and the degree to which, timbral brightness contributes to this delineation.

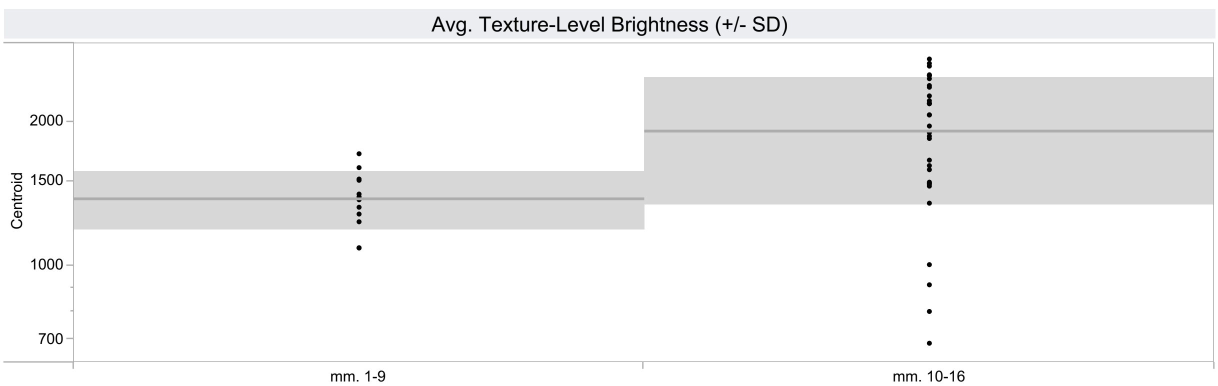


Figure 5-15. Average texture-level brightness of the first two sections. The second section is brighter and more volatile.

Timbre at the level of the texture delineates the form in the sense that each section has a distinct mean brightness and standard deviation. Mean brightness here refers to the overall

[20] This formal segmentation aligns roughly with Morris (1997), although he places the transition a measure sooner (mm. 16) and divides the antiphonal contrast section into two sections (mm. 19–21, and mm. 22–24). Further, Morris has the "closing section" begin with the flute flourish in m. 25, whereas I see that gesture as part of the antiphonal contrast section (although signaling its dissolution).

texture-level brightness averaged over the duration of the formal section in question. The standard deviation, as discussed in Chapter 4, correlates with the section's "volatility" of brightness. This can be gleaned from Figure 5-14a above, but is made clearer in Figure 5-15, which shows two adjacent sections delineated by mean brightness and volatility. While the first section (mm. 1-9) features a mean brightness of approximately 1400 Hz and low volatility score, the next section (mm. 10–16) is decidedly brighter (1900 Hz), and much more volatile.

The layer-level analysis (Figure 5-14c) shows a difference in the textural organization of adjacent sections, also contributing to their delineation. Let us examine the three adjacent textures shown with rectangles in Figure 5-14c. The first section features three solo layers that are widely dispersed in spectral space and thus perceptually segregated. The next section (mm. 10–16), by contrast, begins with a tightly packed texture containing four layers (m. 10), then moves through a series of widely spaced textures involving an ever-growing number of instruments, and finally returns to a tightly packed pair of layers in m. 16. This section also distinguishes itself from its surroundings via the "layer convergence" procedure described previously (not shown in 5-14c), wherein layers become integrated over time. The transitional section features a very different kind of texture than either of the previous sections: we now have just two layers, but they are dense (i.e., containing three instruments), and positioned very close to each other in terms of brightness. We've progressed from dispersed and thin (m. 6) to clustered and dense (m. 18).

The element-level analysis (5-14d), too, shows differences that aid in delineating formal sections from each other. Particularly salient is the difference between the beginning of the "Lourd et sauvage" section (mm. 19–22) and the closing section before the terminal oboe solo (mm. 26–29), both highlighted in rectangles. The element-level centroids in the "Lourd et sauvage" section are on average much closer together than those in the closing section. Thus, there is a clear expansion outward of timbres in the closing section, suggesting an increasing transparency of instruments and greater intralayer heterogeneity.

Beyond emphasizing the main formal divisions of the work, timbre subtly shapes our temporal understanding of the piece in other ways. The large-scale texture-level analysis in Figure 5-14a suggests dividing the piece into two halves, with the extreme dark point at measure 16 marking the separation. There is a timbral symmetry on this reading: the piece begins and ends with a bright oboe solo and is divided at its midpoint by a dark bassoon solo. This perspective reveals a large-scale "bright-dark-bright" timbral gesture across the movement, an overarching gesture mirrored by several small-scale "bright-dark-bright" gestures throughout the movement. These are indicated by dashed brackets in Figure 5-14a. At the smallest scale, even the first three notes of the movement (G♭5–F4–E5 in the oboe) can be considered a "bright-dark-bright" gesture (see the note-level centroids in Figure 5-1b). While many authors, such as MacDonald (2003) and Moura (2004), have noted the "germinal" or "generative" role of this opening oboe solo in terms of its pitch material, my analysis suggests that its timbral shape, too, may serve as a microcosm of the entire movement.

This chapter examines the role of timbre in the first movement of *Octandre* by analyzing brightness differences across multiple levels, from individual elements to integrated layers and the overall texture. By focusing on the multilevel timbral structure of formal sections and their form-building relationships, I have intentionally bracketed parameters traditionally invoked in discussions of form—most glaringly, pitch—in favor of a “timbre first” approach. I have thus demonstrated that analysts can uncover meaningful patterns, trends, gestures, and configurations of brightness that carry the structure of the work on multiple levels, without dissecting the inner workings of its harmony, motivic relationships, or rhythmic gestures. Isolating timbre in this way allows us to make sense of Varèse’s treatment of timbre as an “essential” structural agent. Further work will reintegrate these important domains with the timbral findings enumerated here, evaluating their complex interactions and interdependencies.

More broadly, this chapter has demonstrated the utility of a multilevel approach that accounts for timbre’s multiple roles in post-tonal music. Alternating between and integrating various levels of brightness analysis allows us to remain sensitive to multiple analytical orientations and the different timbral insights they afford. This reflects, ultimately, the overarching goal of the dissertation: to demonstrate an adaptable approach to timbre analysis that bridges the gap between score and audio and allows music theorists to produce accessible and sensitive analyses of a vital but complex element of music.

Conclusion

In this dissertation, I set out to capture the complexities of timbre in post-tonal music through the analysis and visualization of brightness. By quantifying relative brightness at multiple textural levels and temporal scales, I reveal its role in shaping what we hear: we use brightness to determine whether instrumental tones cohere or diverge, whether strands of polyphony go together or stand apart, and whether adjacent formal units contrast with each other and by how much. Further, beyond explicating these details through acoustic measurement and visualization, the dissertation directs our attention to aspects of the musical signal that previously eluded our notice, suggesting new and often surprising stories about how the music goes. To conclude, I will summarize some of the analytical benefits demonstrated in the foregoing chapters and consider some areas for further research.

In Chapter 2, my analysis of Schoenberg's "Farben" focused on element-level brightness. Relationships among concurrent element-level centroids capture the overall "color" of the "Farben" chord—its span, unevenness, average brightness, and "stand-out" timbres—as it changes through time. Vertical ordering of the elements within sonorities, while traditionally considered in terms of pitch, can be reframed in terms of relative brightness and draw our attention to previously unnoticed details: even the lowest-pitched elements in a chord can at times be positioned at the "top" of the sonority. The canonic structure of the piece invites an analysis of "brightness contours"—sequences of centroids that make up each polyphonic voice—revealing a set of brightness segments that repeat and permute each other at important formal junctures.

Chapter 3 proposed relative brightness as a way to think about texture in complex post-tonal polyphony. In Crawford's *Music for Small Orchestra,* rhythmic coordination groups elements into layers, which exhibit differing degrees of homogeneity depending on the relative brightness of the grouped elements. In turn, these layers relate to each other in terms of their overall brightness, either clustering together or remaining distinct, depending on layer-level centroid placement. I argued that analyzing centroid distribution allows us to evaluate the overall perceptual character of the texture (e.g., its "opacity" and "transparency") and showed that the first section of *Music for Small Orchestra* enacts a gradual process of increasing textural opacity. While conventional approaches to texture see the proliferation of rhythmically independent layers as indicative of increasing complexity, attending to brightness paints a more nuanced picture, one that takes into account the perceptual grouping of layers.

In Chapter 4, visualizing trends in overall brightness provided a foundation for interpreting the formal and emotional dimensions of Webern's op. 6 no. 2. Comparative analysis of three recordings revealed that texture-level brightness rises and falls with the dramatic formal divisions of the piece, carrying out a non-linear brightening process that becomes more and more volatile as the piece progresses. These audio-based observations opened up new ways of interpreting both the form and Webern's comments about the piece (as depicting the "certainty" of a "catastrophe") that would not be available through score-based analysis alone.

Chapter 5 integrated brightness analysis at all three textural levels in a close reading of Varèse's *Octandre*. Jumping off from Varèse's idea that timbre functions like an "agent of delineation" in his music, I analyzed the timbral structure of each formal unit of the piece, ultimately arguing that each unit is distinguished from its surroundings by brightness relationships at one or more levels. Large-scale trends in brightness reveal different ways of hearing the temporal structure of the piece beyond these formal units, namely as articulated by a large "bright-dark-bright" gesture that is mirrored at different temporal scales.

Future Directions

I will close by considering some further research directions and addressing the limitations of the current approach. I discuss two broad areas for building on this work: integrating my approach with close analysis of other domains (i.e., pitch and other audio features) and applying the approach to repertoires beyond post-tonal music.

Integrating Other Features

As I mentioned at the close of Chapter 5, salient parameters often invoked in discussions of musical structure, notably pitch, are intentionally bracketed for the purposes of this dissertation but can be reintegrated to enrich my analyses. Because the approach I've taken is so clearly bound to musical units defined by a score—discrete notes, chords, beats, measures, etc.—drawing connections to pitch structures seems readily feasible.

I see two potential strategies here, which I call *alignment* and *integration*. Alignment would involve coordinating a brightness analysis with a separate pitch-based analysis, attending to how the two readings work in tandem or in tension with each other to describe the music. Integration, on the other hand, would involve an analysis of one domain significantly influencing the analysis of another. For example, a prior analysis of layer-level clustering might strongly influence how we segment a complex musical surface into pitch-class sets. Conversely, decisions about grouping elements into layers could be guided by a prior harmonic analysis. In short, while existing analysis of this repertoire has privileged pitch, the timbral alternative I propose need not stand alone, unrelated to the dominant trend. Instead, these approaches should be encouraged to enrich each other and thereby bring about new ways of engaging with and understanding this repertoire.

The approach might be improved by considering other timbral features beyond brightness. My strategy here would be to keep brightness as the central thread of the analysis, while considering additional acoustical information. At the level of the texture, for example, features describing perceived noisiness and inharmonicity might reveal more about how adjacent formal units relate to each other beyond their centroids.[1] At the lower levels, some features that I describe using centroids might be better explained by the features of the overall spectrum at a higher textural level. For example, the "spread" of element-level centroids (defined in Chapter 2 as the SCQ between the lowest and highest centroid) might

[1] For an accessible description of many perceptually relevant audio features and tools for their extraction, see Devaney (2022).

also be explained by *spectral spread* (the standard deviation of frequencies around the centroid) of the *layer* or *texture*.

As described in Chapter 1, I collapse fine-scale temporal fluctuations of the centroid into a single value representing average brightness. Differences between the attack and the steady-state portion of the sound, for example, are not accounted for. This seems somewhat justified by the simple instrumental tones that make up the music analyzed, but I can imagine numerous situations—percussion-dominated music, for example—where analyzing the brightness of a sound's attack and sustain portion separately would be crucial for understanding its effect in the music. Along the same lines, a sound's *attack time*—the time it takes to reach a peak of loudness at the onset of the sound—has been shown to be an important attribute for distinguishing sounds (e.g., Grey and Gordon, 1978). Incorporating this into a brightness-oriented analysis, revealing how sounds behave at their onsets, could enrich our approach by introducing an additional factor contributing to the perceptual "distance" between sounds. This would be particularly useful at the element level and on small temporal scales, when claims about similarity and homogeneity are central.

Applicability to Other Repertoire

While I focus on a narrow selection of pieces in this dissertation, brightness analysis can be applied to a wide range of music, including (but not limited to) both earlier and later Western art music, and recorded popular music. However, the refinement of my approach through close study of a limited repertoire makes some components somewhat less adaptable. Opportunities and potential problems associated with the expansion of analytical scope are addressed below.

Common-practice Western art music seems a promising area for multilevel brightness analysis. The role of orchestration in eighteenth- and nineteenth-century symphonic works, for example, can in part be captured via brightness relationships on multiple textural levels and temporal scales. One potential line of inquiry would be to examine relationships between sound and structure, both in individual works and in corpora: how are pairs of formal units (e.g., first and second theme; exposition and recapitulation) related in terms of brightness? Is there a connection between certain formal functions and texture-level brightness trajectories? Answering these kinds of questions via multilevel brightness would complement existing analyses of orchestration in Western art music that foreground instrumentation (e.g., Dolan, 2013). The general availability of recordings, as well as encoded MIDI data, offers a readily accessible foundation for this work.

One potential complication here arises from evolving instrument technologies through history. Between the eighteenth century and now, most orchestral instruments have undergone drastic mechanical enhancements that alter their timbres, and as such, eighteenth-century compositions have been treated to a wide range of timbral realizations over time. The issue becomes more pronounced with earlier music. This is quite a different situation than the post-tonal music I analyze in this dissertation, which is performed today on largely the same instruments its composers expected to hear. I do not view this as a prohibitive complication, but rather a productive one. It seems to me that analyzing orchestration in eighteenth-century music via the same method described in this dissertation—using modern recordings and orchestral sample libraries—offers valuable insights but represents a slight shift in emphasis away from the timbral decisions of the composer and toward the effects of modern performance. An alternative would be to analyze historically informed performances of the works (e.g., on period instruments) in tandem with orchestral sample libraries of historic instruments (e.g., VSL's "Historic Winds" library). Employing both methods together, in turn, would allow us to address differences and similarities between historic and modern brightness, and make new connections between sound worlds of the past and present.

Vocal music is excluded from consideration in the dissertation, partly owing to the special treatment it requires. Firstly, music featuring voices will often exhibit greater timbral variance across recordings than instrumental music, because singers' voices are so dissimilar.[2] Further, in vocal music the spectral centroid will always strongly correlate to vowel formants and will always spike with certain consonants. This, again, might be interesting to explore deeply in future work—do composers pair bright vowels with bright instruments to create timbral continuity? Do significant words of the poetry have brightness profiles that are reflected elsewhere in the music?[3]

---

[2] For an empirical study that uses audio features (including spectral centroid) to model both inter- and intra-singer similarity, see Devaney (2016).

[3] Cogan's (1969) analyses connecting vowel sounds ("verbal timbre") and melodic line in vocal music suggests one possible path for this future work.

The multilevel approach can also be applied to later twentieth-century music and beyond, but a set of limitations arises from its dependence on discrete tones in a musical score. Noble & McAdams (2020) make the distinction between "post-tonal" music, which explores new combinations of "tones," and "post-tone" music, which explores sonic resources other than tones. Insofar as it is takes elements (i.e., "tones") to be the fundamental building blocks of the music, the multilevel approach seems more easily applied to post-tonal music than "post-tone" music. Its emphasis on pairwise relationships between entities (i.e., SCQs), too, makes it somewhat less apt for describing the continuous sonic processes so common in twentieth-century music.

That said, much experimental instrumental (i.e., "post-tone") music can nevertheless be analyzed via multilevel brightness. Element- and layer-level analysis may require sample libraries that feature extended techniques (e.g., IRCAM's "Solo Instruments"). Alternatively, the sonic and notational complexity of some music may necessitate alternatives to sample libraries: although labor intensive, it would be incredibly fruitful to produce multitrack recordings, which would allow us to isolate elements and layers from the polyphonic totality. This, in turn, would open the approach to a wide range of music without a conventional score.

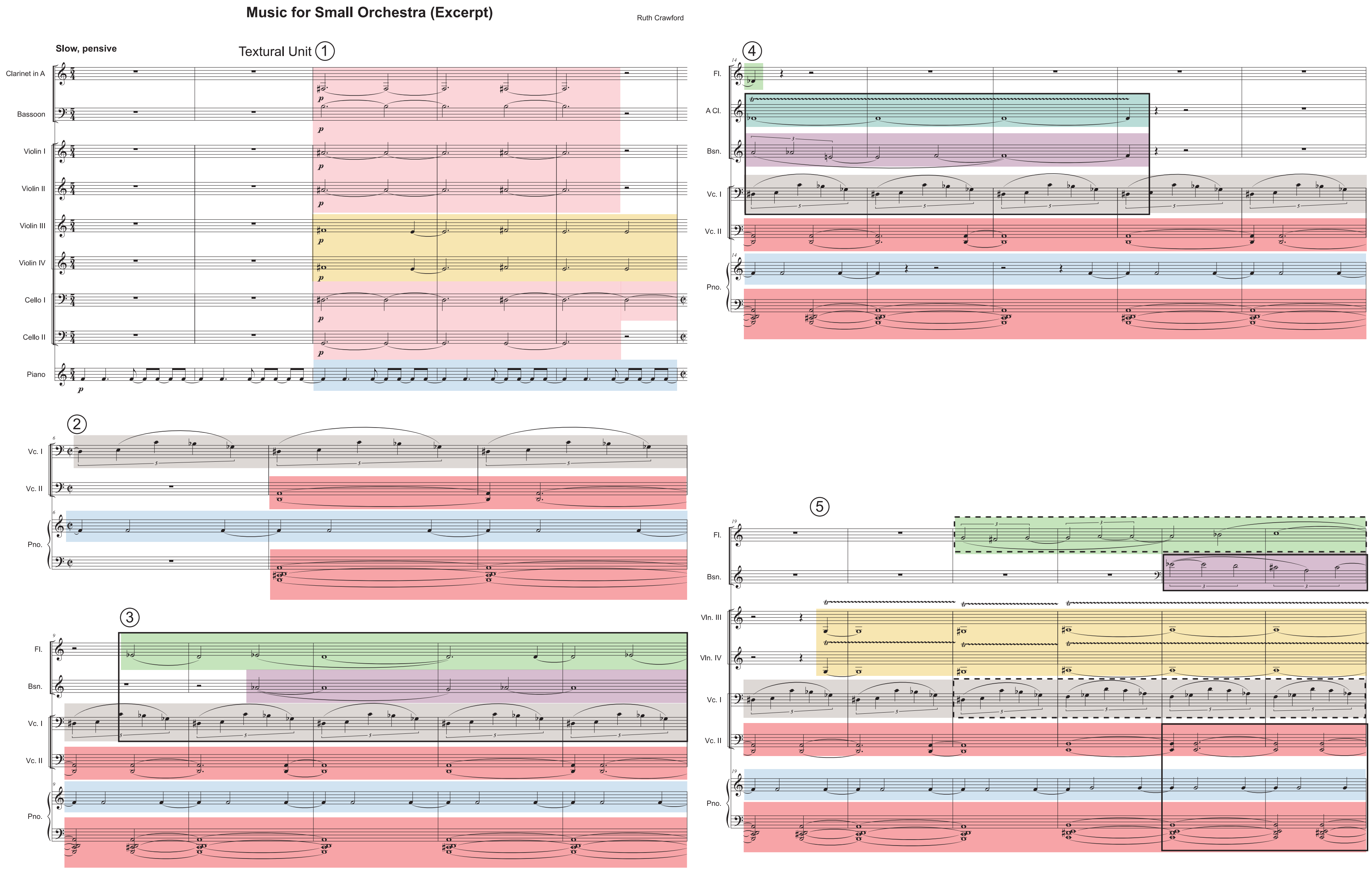
In recent years, there has been a surge of timbre-oriented scholarship surrounding popular music. Multilevel brightness analysis seems particularly applicable to this repertoire. To analyze a song's elements and layers, the most common and viable approach is to analyze "stems" of the song: audio files containing one or multiple sound sources in the recording.

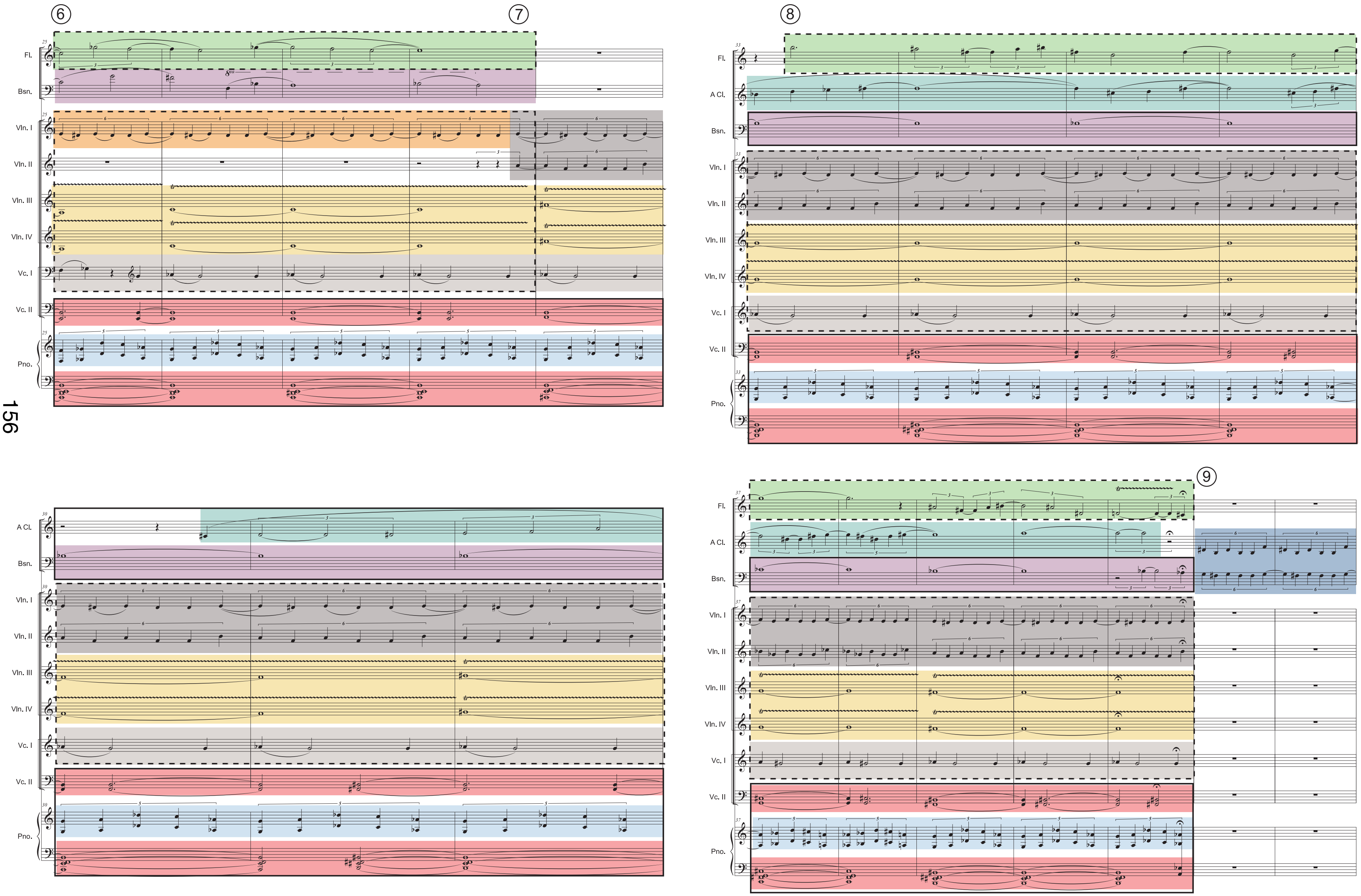
While the original stems of commercial recordings are difficult to obtain legally, several source separation algorithms are currently available that can separate popular music recordings into distinct layers. At the time of this writing most such algorithms separate a recording into four stems: voices, drums, bass, and a stem containing all other sources. Depending on the number of actual sources in the recording analyzed, this breakdown may or may not provide enough separation to do the kind of multilevel brightness analysis demonstrated in this dissertation. Nevertheless, analyzing relative brightness of these separate stems in isolation, in combination, or in dialogue with texture-level brightness at different temporal scales suggests promising directions for timbre-oriented popular music analysis.

In summary, this dissertation offers a highly adaptable approach to analyzing timbre at multiple levels, allowing for flexible and accessible readings of a crucial, yet often daunting, element of music. The method can be effectively combined with analysis of other domains and applied to multiple repertoires, allowing us to shed light on timbre's contribution to the musical structures we hear, and challenge the ways we listen.

Appendix

Annotated score of Ruth Crawford's *Music for Small Orchestra,* mm. 1–43 (refer to Figure 3-10).
Grouped elements shown in same color. Clustered layers encased in same rectangle type (dotted and solid).




6
7
8
9
Fl.
A Cl.
Bsn.
Vln. I
Vln. II
Vln. III
Vln. IV
Vc. I
Vc. II
Pno.

References

Alluri, V., & Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception, 27* (1), 223–241.

Almeida, A., Schubert, E., Smith, J., & Wolfe, J. (2017). Brightness scaling of periodic tones. *Attention, Perception, & Psychophysics*, *79*(7), 1892–1896.

Ancell, J. E. (1960). Sound pressure spectra of a muted cornet. *Journal of the Acoustical Society of America, 32*(1), 1101–1104.

Anderson, J. (1984). *The influence of scientific concepts on the music and thought of Edgard Varèse.* [PhD Dissertation]. University of Northern Colorado.

Anderson, J. (1991). Varèse and the lyricism of the new physics. *The Musical Quarterly, 75*(1), 31–49.

Andrews, G. (1984). *The theory of partitions.* Cambridge University Press.

Babbitt, M. (1966). Edgard Varèse, a few observations about his music. *Perspectives of New Music*, *4*(1), 14–22.

Baker, J. M. (1982). Coherence in Webern's *Six Pieces for Orchestra* Op. 6. *Music Theory Spectrum 4*(1), 1–27.

Berry, W. (1976). *Structural functions in music.* Prentice-Hall.

Bernard, J. (1987). *The music of Edgard Varèse*. Yale University Press.

Bernard, J. (1992). Cracked octaves, warped perspectives: A response. *Perspectives of New Music, 30*(2), 274–289.

Brant, H. (2009). *Textures and timbres*. Carl Fischer.

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound.* MIT Press.

Burkhart, C. (1973). Schoenberg's Farben: An analysis of Op. 16, No. 3. *Perspectives of New Music*, *12*(1/2), 141–172.

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America, 118*(1), 471–482.

Celestini, F. (2009). Der Schrei und die Musik: Mahlers Klänge in Weberns Orchesterstück op. 6 nr. 2. In D. Schweiger and N. Urbanek (Eds.), *Webern_21* (pp. 55–72). Böhlau Wien.

Chou, W. (1966). Varèse: A sketch of the man and his music. *Musical Quarterly, 52*(2), 151–170.

Chou, W. (1978). *Ionisation*: The function of timbre in its formal and temporal organization. In S. Van Solkema (Ed.), *The New Worlds of Edgard Varèse: A Symposium* (pp. 27–74). Institute for Studies in American Music.

Chou, W. (2006). Converging lives: Sixteen years with Varèse. In F. Meyer & H. Zimmermann (Eds.), *Edgard Varèse: Composer, Sound Sculptor,* Visionary (pp. 348–360). Woodbridge: Boydell.

Cogan, R. (1969). Toward a theory of timbre: Verbal timbre and musical line in Purcell, Sessions, and Stravinsky. *Perspectives of New Music, 8*(1), 75–81.

Cogan, R. (1984). *New images of musical sound*. Harvard University Press.

Cogan, R. and Escot, P. (1976). *Sonic design: The nature of sound and music.* Prentice Hall.

Cousineau, M., Carcagno, S., Demany, L., & Pressnitzer, D. (2014). What is a melody? On the relationship between pitch and brightness of timbre. *Frontiers in Systems Neuroscience, 7*.

Cowell, H. (1930/1997). *New Musical Resources.* Cambridge University Press.

Devaney, J. (2016). Inter- versus intra-singer similarity and variation in vocal performances. *Journal of New Music Research, 45*(3), 252–264.

Devaney, J. (2022). Digital audio processing tools for music corpus studies. In D. Shanahan, J.A. Burgoyne, & I. Quinn (Eds.), *Oxford Handbook of Music and Corpus Studies.* Oxford University Press.

Di Gasbarro, F. (2018). The spaced chromatic circle: Varèse's open harmonic system in a nutshell. *Mitteilungen der Paul Sacher Siftung, 31*(1), 24–32.

Dolan, E. (2013). *The orchestral revolution: Haydn and the technologies of timbre*. Oxford University Press.

Duane, B. (2013). Auditory streaming cues in eighteenth- and early nineteenth-century string quartets: A corpus-based study. *Music Perception 13*(1), 46–58.

Duane, B. (2017). Thematic and non-thematic textures in Schubert's three-key expositions. *Music Theory Spectrum, 39*(1), 36–65.

Erickson, R. (1975). *Sound structure in music*. University of California Press.

Falck, R. (1993). Anton Webern's *Six Pieces for Orchestra:* A Comparison of the Two Published Versions. *Canadian University Music Review, 13*(1), 104–122.

Fischer, M., Soden, K., Thoret, E., Montrey, M., & McAdams, S. (2021). Instrument timbre enhances perceptual segregation in orchestral music. *Music Perception, 38*(5), 473–498.

Gentil-Nunes, P. (2013). Partitional analysis and rhythmic partitioning: Mediations between rhythm and texture. *Muzikos Komponavimo Principai/Principles of Music Composing*, *13*(1), 44–51.

Gentil-Nunes, P. (2017). Nesting and intersections between partitional complexes. *MusMat: Brazilian Journal of Music and Mathematics*, *2*(1), 93–108.

Gentil-Nunes, P. (2020). Reading textural functions, instrumental techniques, and space through partition complexes. *MusMat: Brazilian Journal of Music and Mathematics*, *4*(1), 80–97.

Goodchild, M., & McAdams, S. (2018). Perceptual processes in orchestration. In E. Dolan & A. Rehding (Eds.), *Oxford Handbook of Timbre*. Oxford University Press.

Gregory, A. H. (1994). Timbre and auditory streaming. *Music Perception*, *12*(2), Article 2.

Grey, J. M. (1975). *An exploration of musical timbre* [PhD Dissertation]. Stanford University.

Grey, J. M., and Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America*, *63*(1), 1493.

Guck, M. (1992). Varèse bound. *Perspectives of New Music, 30*(2), 244–273.

Guck, M. (1993). The "endless round". *Perspectives of New Music, 31*(1), 306–314.

Guigue, D., and de Paiva, C. (2018). The structural function of musical texture: Towards a computer-assisted analysis of orchestration. *Journées d'Informatique Musicale* (JIM 2018), Amiens, France.

Hardman, K. (2022). The continua of sound qualities for Tanya Tagaq's katajjaq sounds. In L. Shuster, S. Mukherji, & N. Dinnerstein (Eds.), *Trends in World Music Analysis* (pp. 85–99). Routledge.

Hartmann W., and Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception 9*(2), 155–183.

Hasty, C. (1984). Phrase formation in post-tonal music. *Journal of Music Theory, 28*(2), 167–190.

Helmholtz, H. (1863/1875). *On the sensations of tone as a physiological basis for the theory of music* (A. J. Ellis, Trans.). Longman.

Howland, P. (2015). Formal structures in post-tonal music. *Music Theory Spectrum, 37*(1), 71–97.

Huron, D. (1989). Characterizing musical textures. Paper presented at the International Computer Music Conference, San Francisco.

Julius, R. (1978). Edgard Varèse: An oral history project, some preliminary conclusions. *Current Musicology, 25*(1), 38–49.

Kazazis, S., Depalle, P., & McAdams, S. (2021). Ordinal scaling of timbre-related spectral audio descriptors. *Journal of the Acoustical Society of America, 149*(1), 3785–3796.

Kazazis, S., Depalle, P., & McAdams, S. (2022). Interval and Ratio Scaling of Spectral Audio Descriptors. *Frontiers in Psychology*, *13*, 835401.

Kendall, R. A., & Carterette, E. C. (1996) Difference thresholds for timbre related to spectral centroid. In B. Pennycook & E. Costa-Giomi (Eds.), *Proceedings of the Fourth International Conference on Music Perception and Cognition* (pp. 91-95). Montreal: McGill University.

Krimphoff, J., McAdams, S., & Winsberg, S. (1994). Caractérisation du timbre des sons complexes. *Journal de Physique* IV, 4(C5), 625–628.

Krumhansl, C. (1989). Why is musical timbre so hard to understand? In S. Nielzén & O. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43–53). Excerpta Medica.

Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics, 62*(1), 1426–1439.

Lalitte, P. (2011). The theories of Helmholtz in the work of Varèse. *Contemporary Music Review, 30*(5), 327–342.

Lavengood, M. (2020). The cultural significance of timbre analysis: A case study in 1980s pop music, texture, and narrative. *Music Theory Online*, *26*(3).

Lichte, W. H. (1941). Attributes of complex tones. *Journal of Experimental Psychology*, *28*(1), 455–480.

MacCallum, J., & Einbond, A. (2008). Real-Time analysis of sensory dissonance. In R. Kronland-Martinet, S. Ystad, & K. Jensen (Eds.), *Computer Music Modeling and Retrieval. Sense of Sounds* (Vol. 4969, pp. 203–211). Springer Berlin Heidelberg.

MacDonald, M. (2002). *Varèse: Astronomer in sound.* Kahn & Averill.

Maler, A. (2022). Listening to phrase structure and formal function in post-tonal music. *Integral, 35*(1), 45-68.

McAdams, S., Goodchild, M., & Soden, K. (2021). A taxonomy of orchestral grouping effects derived from principles of auditory perception. *Music Theory Online, 28*(3).

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychological Research-Psychologische Forschung. 58*(3), 177–192.

McFee, B., Lostanlen, V., McVicar, M., Metsai, A., Balke, S., et al. (2020). LibROSA/LibROSA: 0.7.2. (2020). Available online: https://librosa.org

Moldenhauer, H. (1979). *Anton Von Webern: A chronicle of his life and work.* Alfred A. Knopf.

Moore, A. (2012). *Song means: Analysing and interpreting recorded popular song*. Ashgate

Moreira, D. (2019). *Textural design: A compositional theory for the organization of musical texture* [PhD Dissertation, Universidade Federal do Rio De Janeiro].

Morris, R. (1997). K, Kh, and Beyond. In J. Baker, D. Beach & J. Bernard (Eds.), *Music Theory in Concept and Practice* (pp. 275–308). Rochester: University of Rochester Press.

Moura, E. (2004). Interaction between the generative cell and symmetrical operations in Varèse's *Octandre. Contemporary Music Review, 23*(1), 17–44.

Noble, J. & McAdams, S. (2020). Sound mass, auditory perception, and "post-tone" music. *Journal of New Music Research, 49*(3), 231–251.

Osborn, B. (2013). Subverting the verse-chorus paradigm: Terminally climactic forms in recent rock music. *Music Theory Spectrum, 35*(1), 23–47.

Plomp, R. (1970). Timbre as a multidimensional attribute of complex tones. In R. Plomp & G. F. Smoorenburg, *Frequency Analysis and Periodicity Detection in Hearing* (pp. 397–414). Suithoff.

Rahn, J. (1980). *Basic atonal theory.* Schirmer.

Reynolds, R. (2013). The last word is imagination: A study of the spatial aspects of Varèse's work (Part I: Written evidence). *Perspectives of New Music, 51*(1), 196–255.

Rousseeuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*(1), 53-65.

Saitis, C., & Siedenburg, K. (2020). Brightness perception for musical instrument sounds: Relation to timbre dissimilarity and source-cause categories. *The Journal of the Acoustical Society of America*, *148*(4), 2256–2266.

Sandell, G. (1991). *Concurrent timbres in orchestration: A perceptual study of factors determining 'blend'.* [PhD Dissertation]. Northwestern University.

Sandell, G. (1995). Roles for spectral centroid and other factors in determining "blended" instrument pairings in orchestration. *Music Perception*, *13*(2), Article 2.

Sandell, G., & Chronopoulos, M. (1996). Identifying musical instruments from multiple versus single notes. *The Journal of the Acoustical Society of America, 100*(4), 2752.

Schoenberg, A. (1911). *Harmonielehre.* Universal Edition.

Schubert, E., & Wolfe, J. (2006). Does timbral brightness scale with frequency and spectral centroid. *Acta Acustica United with Acustica*.

Seashore, C. (1938). *Psychology of Music.* McGraw Hill.

Siedenburg, K., & McAdams, S. (2017). Four distinctions for the auditory "wastebasket" of timbre. *Frontiers in Psychology, 8*(1747).

Siedenburg, K. (2018). Timbral Shepard-illusion reveals ambiguity and context sensitivity of brightness perception. *The Journal of the Acoustical Society of America*, *143*(2), EL93–EL98.

Siedenburg, K., Barg, F. M., & Schepker, H. (2021). Adaptive auditory brightness perception. *Scientific Reports*, *11*(1), 21456.

Singh P.G., and Bregman A.S. (1997). The influence of different timbre attributes on the perceptual segregation on complex-tone sequences. *Journal of the Acoustical Society of America 102*(4), 1943–1952.

Smoot, R. (1986). *The synthesis and manipulation of fused ensemble timbres and sound masses by means of digital signal processing.* [PhD Dissertation]. Ohio State University.

Soden, C. (2020). *Orchestral combinations and transformations in operatic and symphonic music.* [PhD Dissertation]. McGill University.

Straus, J. (2022, July 10). *Analyzing Stravinsky's* Rite of Spring: *Intro to Part One* [Video]. YouTube. https://www.youtube.com/watch?v=AgC7mdZt9tg

Strawn, J. M. (1978). The *Integrales* of Edgard Varèse: Space, mass, element, form. *Perspectives of New Music, 17*(1), 138–160.

Swift, R. (1975). Review of *Sound Structure in Music* by Robert Erickson. *Perspectives of New Music, 14*(1), 148–158.

Tardieu, D., & McAdams, S. (2012). Perception of dyads of impulsive and sustained instrument sounds. *Music Perception, 30*(2), 117–128.

Temko, P., & Spencer, P. (1994). *Practical approach to the study of form in music.* Waveland.

Tsang, L. (2002). Towards a theory of timbre for music analysis. *Musicae Scientiae 6*(1), 23–52.

Varèse, E. & Chou, W. (1966). The liberation of sound. *Perspectives of New Music, 5*(1), 11–19.

Wessel, D. (1979). Timbre space as a musical control structure. *Computer Music Journal, 3*(2), 45–52.

Zacharakis, A., Pastiadis, K., & Reiss, J. (2014). An interlanguage study of musical timbre: Semantic dimensions and their acoustic correlates. *Music Perception*, *31*, 339–358.

Zeller, M. (2023). *Klangfarbenmelodie,* chromophony, and timbral function in Arnold Schoenberg's "Farben". *Music Theory Online, 29*(3).