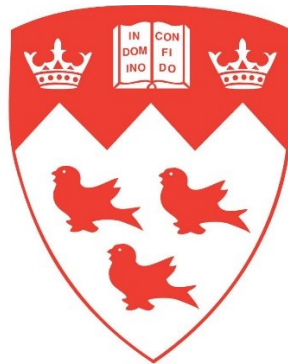


Perceptual interactions of pitch and timbre: An experimental study on pitch-interval recognition with analytical applications

SARAH GATES

Music Theory Area
Department of Music Research
Schulich School of Music

McGill University
Montréal • Quebec • Canada



August 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts.

Copyright © 2015 • Sarah Gates

Contents

List of Figures	v
List of Tables	vi
List of Examples	vii
Abstract	ix
Résumé	xi
Acknowledgements	xiii
Author Contributions	xiv
Introduction	1
Pitch, Timbre and their Interaction • <i>Klangfarbenmelodie</i> • Goals of the Current Project	
1 Literature Review	7
Pitch-Timbre Interactions • Unanswered Questions • Resulting Goals and Hypotheses • Pitch-Interval Recognition	
2 Experimental Investigation	19
2.1 Aims and Hypotheses of Current Experiment	19
2.2 Experiment 1: Timbre Selection on the Basis of Dissimilarity	20
A. Rationale	20
B. Methods	21
Participants • Stimuli • Apparatus • Procedure	
C. Results	23
2.3 Experiment 2: Interval Identification	26
A. Rationale	26

B. Method	26
Participants • Stimuli • Apparatus • Procedure • Evaluation of Trials • Speech Errors and Evaluation Method	
C. Results	37
Accuracy • Response Time	
D. Discussion	51
2.4 Conclusions and Future Directions	55
3 Theoretical Investigation	58
3.1 Introduction	58
3.2 Auditory Scene Analysis	59
3.3 Carter Duets and <i>Klangfarbenmelodie</i>	62
<i>Esprit Rude/Esprit Doux</i> • Carter and <i>Klangfarbenmelodie</i> : Examples with Timbral Dissimilarity • Conclusions about Carter	
3.4 Webern and <i>Klangfarbenmelodie</i> in Quartet op. 22 and Concerto op 24	83
Quartet op. 22 • <i>Klangfarbenmelodie</i> in Webern's Concerto op. 24, mvt II: Timbre's effect on Motivic and Formal Boundaries	
3.5 Closing Remarks	110
4 Conclusions and Future Directions	112
Appendix	117
A.1,3,5,7,9,11,13 Confusion Matrices for each Timbre Pair	
A.2,4,6,8,10,12,14 Confusion Matrices by Direction for each Timbre Pair	
B.1 Response Times for Unisons by Timbre Pair	
References	122

List of Figures

Fig. 2.1	3D Timbre Space	25
Fig. 2.2	Screen Shot of Evaluation Program	34
Fig. 2.3	Response Accuracy as a Function of Interval for the Two Interval Directions	40
Fig. 2.4	Mean Accuracy (%) by order (FH-PN)	41
Fig. 2.5	Root-mean squared errors (in semitones) for each timbre pair	44
Fig. 2.6	Response times (in seconds) for the identification of different intervals for each Timbre Pair	47
Fig. 2.7	Response times (in seconds) for the identification of different intervals for each direction	48
Fig. 2.8	Response Times (seconds) for Order, pair FH-PN	50
Fig. 2.9	Response Times for Intervals by Order for MB-VN	51
Fig. 3	Timbre Space, similarity of flute and clarinet	65
Fig. 4.	3D timbre space, winds vs. strings	75

List of Tables

Table 2.1	Timbre Toolbox Descriptors and Correlation Coefficients	25
Table 2.2	Timbre Pairs and 3D Space Coordinates Differences	26
Table 2.3	Interval Distributions for interval experiment	29
Table 2.4	Pitch Distributions for interval experiment	30
Table 2.5	Mean Accuracy for Unisons by Timbre Pair	38
Table 2.6	Mean Accuracy for Timbre Pair (all other intervals 1-12)	39
Table 2.7	Mean Accuracy by Direction	39
Table 2.8	Mean Accuracy by Interval	39
Table 2.9	Root Mean Square values for each timbre pair	44
Table 2.10	Response Times for Timbre Pair (all other intervals 1-12)	46
Table 2.11	Response Times for Direction	46
Table 2.12	Response Times for each Interval (LogRT left, RT in seconds, right)	47
Table 2.13	T-test results on response times for intervals by direction	48
Table 2.14	Mean Response Times by Order for FH-PN (LogRT's)	49
Table 2.15	Mean Response Times by Order for MB-VN	50
Table 2.16	Mean Response Times by Order for BC-MT	50
Table 2.17	Mean Response Times in seconds for all timbre pairs	50

List of Examples

Example 1.1	<i>Esprit Rude/Esprit Doux</i> , mm 1-4	66
Example 1.2	<i>Esprit Rude/Esprit Doux</i> , mm 1-4, combined version	67
Example 1.3	<i>Esprit Rude/Esprit Doux</i> , mm 1-4, phrase divisions	68
Example 2.1	Carter <i>Esprit Rude/Esprit Doux</i> , mm 83-88	71
Example 2.2	<i>Esprit Rude/Esprit Doux</i> , mm 83-88, combined version with phrase boundaries	73
Example 3	<i>Esprit Rude/Esprit Doux</i> , mm 32-35	74
Example 4.1	<i>Rigmarole</i> , mm 15-25	76
Example 4.2	<i>Rigmarole</i> , mm 15-25, piano (timbre neutral)	79
Example 5.1	<i>Au Quai</i> , mm 12-17, perceived monophonic line	80
Example 5.2	<i>Au Quai</i> mm 12-17, piano (timbre neutral) phrase boundaries	82
Example 6.1	Quartet op. 22, mvt II, mm 68-88, original	85
Example 6.2	Quartet op. 22, mm 75-78, monophonic line	86
Example 6.3	Quartet op. 22, measures 81-84, monophonic line	87
Example 6.4	Quartet op. 22, mm 70-72	90
Example 6.5	Quartet op. 22, measures 85-88	90
Example 6.6	Quartet op. 22, mm 68-88, summary	92
Example 7.1	Concerto op. 24, mvt II, mm 1-28, analysis summary of Bailey, Wintle and Spinner	97
Example 7.2	Concerto op. 24, mvt II, mm 1-28, piano reduction	100

Example 7.3	Concerto op. 24, mvt II, mm 1-11, Antecedent phrase in Clarinet version	102
Example 7.4	Concerto op. 24, mvt II, mm 11-22, Consequent phrase in Clarinet version	103
Example 7.5	Concerto op. 24, mvt II, mm 23-28, Extension in Clarinet version	103
Example 7.6	Concerto op. 24, mvt II, mm 1-10, Antecedent phrase constructed as a Sentence	105
Example 7.7	Concerto op. 24, mvt II, mm 11-22, Consequent phrase constructed as Sentence	108
Example 7.8	Concerto op. 24, mvt II, mm 23-28, Extension	109

Abstract

The goal of this project is to investigate pitch-timbre interactions using a two-pronged approach. The first part is an experimental study to examine how change along different dimensions of timbre affects musicians' ability to categorically identify pitch intervals, while the second part is a music-theoretic investigation of the effects of pitch and timbre on the perception of *Klangfarbenmelodie* in the works of Carter and Webern, using concepts from auditory scene analysis developed by Albert Bregman and others. The primary objective is to observe pitch-timbre interactions in both experimental and music-theoretic settings in order to better understand the role of such interactions in music perception.

In the experimental study, musicians with relative pitch ($n=22$) identified melodic intervals in ascending and descending directions within the octave in both timbre-neutral (piano only) and timbre-changing conditions. Multidimensional scaling analysis of timbre similarity ratings was used to select timbre pairs that varied along both spectral dimensions (e.g., spectral centroid) and temporal dimensions (e.g., effective duration and amplitude modulation). Contrary to the primary hypothesis, the intervals that were poorly identified in the timbre-neutral conditions were not more susceptible to interference with timbre, suggesting that the role of musical training in interactions is more complicated than previously thought. Some evidence was found supporting the hypothesis that spectral features of timbre interfere with pitch perception, although the overall effect of timbre change was inconclusive as the timbre-neutral condition did not outperform any timbre-changing condition in either accuracy or response time. The results indicate that more research is needed to understand the interactions of pitch and timbre, and that

these interactions are likely highly dependent on the experimental task and the participant population and skillset.

In the music-theoretic investigation, the principles of auditory scene analysis are applied to analyses of various *Klangfarbenmelodien* of both Elliott Carter and Anton Webern. The roles of both pitch and timbre are discussed in order to define what makes possible the perception of an unbroken, sequentially integrated Webernian *Klangfarbenmelodie*. Analyses of the original versions and timbre-neutral piano reductions of several duets by Carter (*Esprit Rude/Esprit Doux*, *Rigmarole*, and *Au Quai*) reveal the importance of timbral similarity for the perception of sequentially integrated *Klangfarbenmelodie* in two-instrument settings. Analysis of the second movement of Webern's Quartet, op. 22 extends the work on Carter duets to the discussion of a more complicated, four-voice texture. The effects of pitch and timbre on motivic grouping, phrase boundaries, and formal construction are also discussed in an analysis of the second movement of Webern's Concerto for Nine Instruments, op. 24. I compare my analysis to previous formal analyses of the second movement by Leopold Spinner, Christopher Wintle and Kathryn Bailey to highlight the importance of considering orchestration in analysis.

Résumé

Le but de ce projet est d'examiner les interactions entre le timbre et la hauteur en utilisant une approche à deux volets. Le premier volet est une étude expérimentale qui examine comment le changement le long de différentes dimensions timbrales affecte la capacité des musiciens à identifier catégoriquement les intervalles de hauteur, alors que le second volet est un examen musico-théorique des effets de la hauteur et du timbre sur la perception de *Klangfarbenmelodie* dans les œuvres de Carter et Webern utilisant les concepts de l'analyse de scènes auditives développés par Albert Bregman et d'autres. L'objectif premier est d'observer les interactions entre le timbre et la hauteur à la fois dans des contextes expérimental et théorique à fin de mieux comprendre le rôle de ces interactions dans la perception musicale.

Dans l'étude expérimentale, des musiciens (n=22) ont identifié des intervalles mélodiques, avec hauteur relative, en direction ascendante et descendante à l'intérieur d'une octave dans des conditions de timbre neutre (piano seulement) et de changement de timbre. Une analyse d'échelle multidimensionnelle des évaluations de dissemblance entre timbres a été utilisée pour sélectionner des paires de timbre qui varient le long de dimensions spectrale (le centroïde spectrale) et temporelles (durée effective de l'enveloppe d'amplitude et modulation d'amplitude). Contrairement à l'hypothèse principale, les intervalles mal identifiés dans les conditions de timbre neutre n'étaient pas plus susceptibles d'interférence que ceux ayant des changements de timbre, ce qui suggère que le rôle de la formation musicale dans le domaine des interactions entre le timbre et la hauteur est plus complexe que précédemment supposée. Quelques preuves ont été trouvées soutenant l'hypothèse que les caractéristiques spectrales du

timbre interfèrent avec la hauteur, mais l'effet global de la variation du timbre était relativement peu concluant puisque la condition de neutralité timbrale n'a surpassé aucune condition de changement de timbre tant au niveau de la précision qu'au niveau du temps de réponse. Les résultats indiquent que des recherches supplémentaires sont nécessaires pour comprendre les interactions entre le timbre et la hauteur et que ces interactions sont probablement hautement dépendantes de la tâche expérimentale, la population de participants et leurs compétences musicales.

Dans les investigations musico-théoriques, les principes de l'analyse de scènes auditives sont appliqués à l'analyse de plusieurs *Klangfarbenmelodien* par Elliott Carter et Anton Webern. Les rôles du timbre et de la hauteur sont discutés pour définir ce qui permet la perception d'une *Klangfarbenmelodie* de Webern comme séquentiellement intégrée et non-fragmentée. Des analyses de plusieurs duos originaux par Carter (*Esprit Rude/Esprit Doux*, *Rigmarole*, et *Au Quai*) et de leurs réductions à un timbre neutre au piano révèlent l'importance de la similarité timbrale dans la perception de *Klangfarbenmelodie* séquentiellement intégrées dans le contexte de deux instruments. Une analyse du second mouvement du Quatuor à Cordes, op. 22 de Webern élargit la discussion au sujet des duos de Carter à une discussion plus complexe sur une texture à quatre voix. Les effets de la hauteur et du timbre sur le groupement motivique, les frontières de phrases et les constructions formelles sont aussi discutés dans une analyse du second mouvement du Concerto pour Neuf Instruments, op. 24, de Webern. Je compare mon analyse à des analyses formelles antérieures du second mouvement par Leopold Spinner, Christopher Wintle et Kathryn Bailey afin de souligner l'importance de la prise en compte de l'orchestration dans l'analyse musicale.

Acknowledgements

I would like to thank several people for their support of this project over the last year.

Firstly, I would like to thank my academic advisors Stephen McAdams and Robert Hasegawa, as well as the MPCL Lab Technician Bennett Smith. Stephen McAdams for his help with the experimental design and implementation, use of the MPCL's resources, and for his generous help and instruction regarding data analysis. Robert Hasegawa for his continuous input and support during the experimental phase of the project, incredible knowledge of the repertoire and theoretic literature, as well as for pushing me to make deeper connections between my analytic findings and modern psychological research. Working with both of you has truly been an honour. And Bennett Smith for his incredibly hard work developing the experimental software and stimuli preparation, and particularly for the custom made evaluation program for the intervals experiment. I would not have been able to do this project without your help.

I would also like to thank all of the members of the MPCL for their input as well as for their participation in piloting and loudness matching. A special thanks to Kai Siedenburg for his help with MATLAB during the timbre similarity experiment. To all of the participants of both of my experiments, this would not have been possible without you! For the support of all my friends and family, and to all of my colleagues and professors who have inspired me to learn and grow. And finally, to Guillaume Viger for all of your encouragement and support over the last year. Your support has meant the world to me.

Author Contributions

As the author of this thesis, I was responsible for the research, experimental design, data analysis and interpretation of the results, music-theoretical analyses, and writing and editing of the thesis. My co-adviser, Stephen McAdams provided access to laboratory equipment and resources, as well as guidance on the experimental design, data analysis, and interpretation of the results. My other co-adviser, Robert Hasegawa, provided input on the musical analyses, and guidance on reference materials and repertoire selection. MPCL Technical Manager, Bennett K. Smith, created the evaluation program for the intervals experiment, and also helped with stimulus preparation and implementing the experimental design. Additionally, I was financially supported by the Social Sciences and Humanities Research Council of Canada in the form of a Master's scholarship, as well as an NSERC grant and Canada Research Chair awarded to Stephen McAdams.

Introduction

The role of timbre and orchestration in music-theoretic discourse has been greatly underrepresented in music analysis. The priorities in analysis have been placed instead on the organization of pitch materials, as well as formal, rhythmic, and motivic constructions. This is particularly true of analysis of music of the twentieth century, in which the role of orchestration and timbre has arguably played a more important role than in previous musical styles. Modern psychological and acoustic research has demonstrated, however, that pitch and timbre are not entirely separable acoustic dimensions, and in fact, that they interact in perception (Melara & Marks, 1990a, b, c; Krumhansl & Iverson, 1992; Singh & Hirsh, 1992; Pitt, 1994; Silbert, Townsend, & Lentz, 2009; Caruso & Balaban, 2014; Cousineau, Carcagno, Demany, & Pressnitzer, 2014; Vurma, 2014). What makes these interactions possible? Do these interactions affect music perception in real listening situations? What are the implications for music-theoretic analysis? This project focuses on addressing these questions through a two-pronged approach: firstly, an experimental study to further examine pitch-timbre interactions through investigating the effect of timbre change on interval recognition in musician populations, and secondly, music-theoretic analyses of several twentieth century pieces using the principles of auditory scene analysis in order to understand the role of both pitch and timbre on the perception of *Klangfarbenmelodien*.

Pitch, Timbre, and their Interaction

How is it that pitch and timbre can interact in perception? Pitch and timbre are both perceptual dimensions of a tone, which share some features in common. Pitch is typically

described as being comprised of two dimensions: pitch height and pitch chroma. Pitch height is a tonotopic dimension responsible for contour perception and the sensation of “high” and “low,” which is related to the fundamental frequency, F0, and measured in Hertz. Pitch chroma is more based on temporal fine structure or periodicity, is related to the place of pitch in the musical scale, and is responsible for the sensation of octave equivalence (e.g., C3 vs. C4) (Bigand & Tillmann, 2005). There are two basic types of theories which have been proposed to explain how frequency information is extracted by the auditory system in the perception of complex tones. These are *place theories* and *temporal theories* (Oxenham, 2013). In place theories, it is proposed that the auditory system uses the resolved harmonics¹ to extract the F0 through a template-matching process that associates specific harmonic content to a specific F0 (termed tonotopic organization). Temporal theories generally evaluate time intervals between spikes in the auditory nerves, which involves both resolved and unresolved harmonics (Oxenham, 2013). There is little evidence about which mechanisms are indeed responsible for pitch perception: both place and temporal theories appear to play significant roles. What is clear is that the harmonic spectrum is very important for the perception of complex tones, particularly the resolved harmonics 1-5 (Oxenham, 2013).

Timbre is also multidimensional in nature, although it contains more perceptual dimensions than pitch, and is therefore far more difficult to measure. Traditionally, timbre is defined as “[a] term describing the tonal quality of a sound...[where two tones] sounding the same note at the same loudness are said to produce different timbres.” (Campbell, M, 2015). This negative definition does little to describe what timbre is, and provides no insight into the

¹ Resolved harmonics are the lower end harmonics, generally 1-10, and harmonics 11-24 are unresolved. Resolved harmonics produce excitation patterns on the basilar membrane, which can be perceived independently with cognitive effort under certain conditions (Oxenham 2013).

perceptual features which define it. Timbre has many perceptual dimensions, such as attack sharpness and brightness (McAdams, 2013). Timbral brightness, a perceptual property of spectral centroid (the average of harmonics above the fundamental), is one property that intersects with pitch perception because the spectral features of timbre in particular covary with pitch and dynamics (McAdams, 2013). Because both timbral brightness and pitch height involve assessment of the spectral content of a tone, they are susceptible to interference in perception. This is likely due to the neural representations that share common attributes, such as tonotopic organizations² in the brain (McAdams, 2013). These interactions have been shown to be so robust as to allow for an *interval illusion* (Russo & Thompson, 2005b). Interval illusion is a phenomenon where timbre change along the dimension of brightness causes melodic pitch-intervals to be perceived as larger or smaller than they actually are when relative size is judged.

Klangfarbenmelodie

How might pitch-timbre interactions apply to music-theoretic analysis then? The complex nature of real musical situations create a sincere difficulty for fully understanding the contributing factors of both pitch and timbre on the perception of musical phenomena such as form, phrase structure, motives and pitch organization. The phenomenon of *Klangfarbenmelodie* is a perfect medium for observing these interactions in real musical situations as both pitch and timbre change are of equal importance.

The idea of *Klangfarbenmelodie* is one that originated from Arnold Schoenberg, mentioned at the end of his book *Theory of Harmony* (1948). He refers specifically to “tone

² The mapping of frequency to location in the anatomic organization of the cochlea where sound frequencies are received by specific receptors in the inner ear. This frequency map is preserved at sites in the auditory brain.

colour melodies,” which adhere to a similar logic as applies to pitch melodies (Schoenberg, 1948, pp. 421). This implies that a progression of tone colours should be able to contain and demonstrate a similar perceptible syntax to that which pitch is able to portray. Predating much of the current research on pitch-timbre interactions, Schoenberg even recognizes the complex nature of tone colour and pitch, indicating that pitch is simply a dimension of tone colour, measured in one direction (Schoenberg, 1948). Schoenberg’s *5 Pieces for Orchestra*, op. 16, no 3 “Farben” (1909) is often cited as the exemplar of true *Klangfarbenmelodie* (Dahlhaus, 1987). The definition and realization of (“true”) *Klangfarbenmelodie* is greatly disagreed upon though, as the originating description given by Schoenberg is rather vague. It is currently described often as a melody in which a succession of timbres is at least as important as a succession of pitches (Mathews, 2006), or where a type of balance is achieved between instrumentation and pitch melody (Dahlhaus, 1987).

The lack of clarity in Schoenberg’s definition of *Klangfarbenmelodie* led to differing opinions of what *Klangfarbenmelodie* actually was. One of Schoenberg’s closest students, Anton Webern, took a very different approach to the technique. Therefore, there is often a distinction made between Schoenbergian and Webernian *Klangfarbenmelodien* (Iverson, 2009). Schoenbergian *Klangfarbenmelodie*, as embodied by his 1909 work *5 Pieces for Orchestra*, no 3 “Farben,” is more concerned with vertical blend, focusing on vertical sonorities through use of chords, voice leading and blend. Cramer (2002) suggests that a Schoenbergian *Klangfarbenmelodie* is therefore a harmonic principle, which is why some authors refer to it as *Klangfarbenakkord* rather than melody (Iverson, 2009). Webernian *Klangfarbenmelodie* is more concerned with horizontal, melodic connections, which are more often described as *pointillistic* (Iverson, 2009). This led to the more commonly understood version of *Klangfarbenmelodie*, in

which each pitch in a melodic line is coloured by a different timbre (as in Webern's version of Bach's *Ricercar*), which led to serial techniques of timbre used by the Darmstadt school (Iverson, 2009). Schnittke refers to these varying versions of *Klangfarbenmelodie* as having either timbral consonance (timbral harmony) or timbral dissonance (with timbral counterpoint) (Mathews, 2006). This vertical versus horizontal structure of *Klangfarbenmelodie* is also a key factor in the definition of auditory streams (either sequentially or vertically integrated) in auditory scene analysis (Bregman, 1990). The current work investigates how timbre change can affect the perception of melodic intervals within musician populations.³ The more traditional horizontal *Klangfarbenmelodie* of Webern is therefore an ideal real-world musical phenomenon to investigate in terms of pitch-timbre interactions as the complexities involved with pitch perception and vertical blend have not yet been fully explored by modern psychological research.

Goals of the Current Project

The goals of the current project are to expand the current experimental research on pitch-timbre interactions, while concurrently investigating the possible role pitch-timbre interactions play in the perception of real music. The psychological experiments investigate the effect of timbre on the identification of melodic pitch-intervals within the octave in musician populations. This class of listeners was used specifically to control for and be able to measure interval recognition skills, in order to better understand the role pitch-interval identification plays in the interactions of pitch and timbre. The music-theoretic investigations include detailed analysis of primarily horizontal, Webernian-style *Klangfarbenmelodie* in the music of Elliott Carter and Anton Webern. The

³ For the current study, "musician population" refers to undergraduate and graduate students studying music as their primary major. See participant information, pages 22 and 27.

analysis focuses on the properties of pitch and timbre that allow the perception of a fluid, horizontal flowing *Klangfarbenmelodie* using the principles of auditory scene analysis developed by Albert Bregman and others. The work is separated into two primary sections that (1) overview the experiments completed and results found, and (2) discuss the music-theoretic analyses of Carter and Webern. Connections between the experimental results and theoretic analysis will be made wherever possible, and the difficulties of applying experimental results to music analysis will also be highlighted in a concluding section. I will begin by outlining the research that has been completed on perceptual interactions between pitch and timbre to date, then explaining the experimental design and results.

Literature Review

Pitch-Timbre Interactions

In a pioneering set of studies published in 1990, Robert Melara and Lawrence Marks investigated how different musical dimensions interact in perception, particularly pitch, timbre and loudness. Through several psychological experiments, they proposed a revision to the commonly accepted model of perceptual interactions developed by W. R. Garner in the 1970's. Garner's model classifies perceptual dimensions as either *interacting* or *separable* (Garner, 1974). The Garner model states that stimuli formed by combining attributes on *interacting* dimensions are perceived initially as unidimensional (only one dimension is perceived), but that with cognitive effort, multiple dimensions can be perceived after the initial holistic stage of processing. This differs from stimuli with *separable* dimensions which can be perceived as constituent dimensions immediately, or without cognitive effort. Melara & Marks proposed, however, that contrary to Garner's assertion that interacting dimensions entail an initial absence of dimensionality, perceivers actually have immediate access to dimensions in stimuli that are both interacting and separable (Melara & Marks, 1990c). In their model, perceivers can perceive multiple dimensions in stimuli that have interacting dimensions, but that the dimensions are constrained by one another perceptually since change in one of the dimensions affects the perception of the other. To measure this, a set of tasks, termed Garner classification tasks are used to test how perceivers are able to classify change in one perceptual dimension while the other dimension of the stimuli either changes or stays the same. In their series of experiments, Melara & Marks (1990a, b, c) found that pitch and timbre are *Garner integral*, indicating that the

dimensions of pitch and timbre can be perceived separately, but that change in either one of the dimensions affects the perception of the other. While most of the studies completed after Melara & Marks show perceptual interactions between pitch and timbre, there have been some studies that show they are in fact independent of one another (Semal & Demany, 1991; Marozeau, de Cheveigné, McAdams, & Winsberg, 2003). These differences may be the result of the different tasks used in the studies. Contrary to the Garner classification paradigm used by Melara & Marks, both Semal & Demany (1991) and Marozeau *et al.*, (2003) only required participants to make judgments on one dimension while the other dimension changed. Semal & Demany (1991) required participants to classify pitch alone while timbre changed (labelling two tones in a sequence as either the “same” or “different”), while Marozeau *et al.*, (2003) required participants to make similarity judgments on two timbres while pitch changed. Both of these tasks may have caused participants to complete their ratings without interference because they were not required to consider change in the irrelevant dimension. The majority of the literature to be discussed however, demonstrates that perceptual interactions occur between pitch and timbre, many of which require participants to equally evaluate both dimensions of pitch and timbre.

Following Melara & Marks, several subsequent studies utilized Garner classification tasks to examine pitch-timbre interactions. Several experiments conducted by Krumhansl & Iverson (1992) investigated perceptual interactions between pitch and timbre using speeded classification tasks. Subjects (with a minimum of 5 years of musical training) were asked to classify change in stimuli (in pitch and/or timbre), while either pitch, timbre or both changed. Two synthesized timbres were used (piano and trumpet), as well as two pitches (F4 and C5). Similarly to Melara & Marks (1990b), they found pitch and timbre to be *Garner integral*, and interfered symmetrically with one another. A similar study using speeded classification

conducted by Mark Pitt (1994), used two groups of subjects (musicians and non-musicians) and acoustic timbres from the McGill Masters samples (trumpet and piano), with the pitches 294Hz and 417Hz (D4 and G#4). Subjects were asked to classify if pitch, timbre or both changed or stayed the same. Similarly to Melara & Marks (1990a, b, c), and Krumhansl & Iverson (1992), Pitt found a symmetrical interference for musicians, but unlike the previous studies, found asymmetrical interference for non-musicians, with greater inference in the pitch-focus condition, indicating that timbre change disrupted pitch processing more than the reverse. This result suggests that non-musicians are more affected by timbre change than musicians,⁴ who demonstrated symmetrical interference between the two dimensions.

Other studies have also shown differences between musicians and non-musicians in the interference of pitch and timbre. Beal (1985) conducted similar speeded classification tasks on musicians and non-musicians. Subjects were asked to classify change (pitch, instrument, or none) for both tonal/diatonic chords and atonal chords on acoustic timbres (acoustic guitar, piano, and harpsichord). Beal found that musicians far outperformed non-musicians, particularly in classifying instrument change when pitch remained constant in the diatonic chord condition. Similarly to Pitt's 1994 study, this result suggests that non-musicians are more affected by timbre change than musicians. When the chords were non-diatonic, however, musicians had similar error rates as non-musicians when timbre changed, indicating that when learned diatonic pitch structures were not available for musicians, they performed similarly to non-musicians.

⁴ Generalized claims such as these have been made throughout the paper in order to formulate hypotheses regarding pitch-timbre interactions. The author recognizes that these claims cannot be generalized to larger populations due to the limited amount of research available. With this in mind, references to "musicians" and "nonmusicians" refer to subjects used in previous studies and in the current study (see participant information p. 22 and 27), which are typically undergraduate and graduate students. Issues with limited subject populations in psychological research has been cited elsewhere (see Jones, 2010).

This result is supported by findings in Warrier & Zatorre (2002), who found that tonal context reduced the effect that timbre change had on pitch classification.

Differences between musicians and non-musicians were also found by Russo and Thompson in their 2005 study on pitch-timbre interactions. These interactions were shown to be so strong that they induced an “interval illusion:” where certain melodic pitch intervals could sound larger or smaller than they actually were when timbral brightness changed from one pitch of the interval to the next. Using a subjective rating task (Experiments 1 & 2) and direct comparison task (Experiment 3), they found that melodic tritones with congruent timbre change (timbre change going in the same direction as pitch change, e.g., dull to bright timbre change on an ascending interval) were rated as larger than a perfect fifth containing incongruent timbre change (e.g., bright to dull timbre change on an ascending perfect fifth). This effect was so robust that even musicians labelled congruent tritones as larger than incongruent perfect fifths. Similarly to the study by Mark Pitt (1994), this effect also showed differences between musicians and non-musicians, most notably asymmetric interference in their musician population with symmetric interference in their non-musician population. Non-musicians demonstrated symmetric interval illusion (illusion in both ascending and descending melodic pitch intervals), where musicians demonstrated interval illusion in the descending direction only. The authors suggest an effect of training, claiming that musicians are exposed to fewer large descending intervals than large ascending intervals, thusly making them more susceptible to timbre-induced interval illusion in the descending direction. They support this hypothesis by citing Vos & Troost (1989) who found in a statistical analysis of interval content of classical Western musical sources that there are fewer descending fifths and tritones present than ascending ones.

Not all work on pitch-timbre interactions demonstrate that there are differences between musicians and non-musicians however. Some studies have found that the amount of interference between pitch and timbre is the same for musicians and non-musicians, with the only differences being that musicians are generally more precise in their categorizations of pitch and timbre (Vurma, Raju, & Kuuda, 2010; Zarate, Ritson, & Poeppel, 2013). A recent study by Allen & Oxenham (2014) demonstrated that when differences in sensitivity to the dimensions of pitch and timbre (spectral centroid in particular) are controlled, symmetrical interference is found between pitch and timbre for both musicians and non-musicians.

Much of the previous research reviewed above shows that pitch and timbre interact in perception, and that these interactions are typically symmetrical (Melara & Marks, 1990; Krumhansl & Iverson, 1992; Allen & Oxenham, 2014), although there is some evidence to suggest that pitch and timbre are independent of one another (Semal & Demany, 1991; Marozeau *et al.*, 2003). Within the research demonstrating pitch-timbre interactions, some differences have been shown to exist between musicians and non-musicians, suggesting a training effect for musicians (Russo & Thompson, 2005b, Beal, 1985), and that non-musicians are more influenced by timbre (Pitt, 1994). Many of the authors suggest a training explanation, proposing that the musician's superior pitch-processing abilities (Micheyl *et al.*, 2006) and/or superior analytical listening (Oxenham, Fligor, Mason, & Kidd, 2003) allow them to attend to change in these dimensions more easily than non-musicians. This hypothesis is supported by the finding that when musicians are presented with atonal pitch materials (which typically do not receive attention in ear training), they perform as poorly as non-musicians in classifying pitch change when timbre is varied (Beal, 1985). The access to tonal hierarchical structures might therefore limit timbre's effect on pitch. Warrier & Zatorre (2002) showed that when tonal context is

available to musicians, timbre's influence on pitch judgments is reduced. Some contrasting research however suggests that symmetrical interference is possible between musicians and non-musicians when sensitivity for the different dimensions are controlled (Allen & Oxenham, 2014), and that musical training does not provide a significant advantage in tasks in which both pitch and timbre change (Borchert, Micheyl, & Oxenham, 2011).

Unanswered Questions

There are several issues left unresolved from the previous studies. The contradicting research regarding if and how pitch and timbre interact in perception reveals that the specifics of the phenomena are generally not well understood. Differences found between musicians and non-musicians are also not well understood, which points to a lack of understanding of both the underlying mechanisms responsible for processing timbre and pitch, as well as a lack of understanding of the effects of musical training on the phenomenon. The studies that investigate differences between musicians and non-musicians have not quantified the musician's pitch processing abilities, which is the skill that most researchers hypothesize as being responsible for the differences between musicians and non-musicians. This is a pressing concern, particularly for timbre's effect on the perceived size of intervals investigated by Russo & Thompson (2005b). The authors assume that explicit knowledge of interval categories possessed by their musicians should have prevented timbral interference in pitch interval judgments. The possession of categorical knowledge (and skill of recognition) of interval categories, however, was never verified by the authors. None of the trained musicians in any of the mentioned experiments had their pitch-interval identification skills tested prior to completing the experiment, so many of the musicians (especially those without formal aural skills or theory training, given that mean ages were often around 18 years of age) could have been more like non-musicians in their

identification of pitch intervals.

These differences found between musicians and non-musicians may have also simply resulted from the type of task used. Many of these studies, particularly Russo & Thompson (2005b) merely utilized subjective or comparative rating tasks, which are very different tasks than categorical identification of intervals using interval names (e.g., perfect fifth). Subjective ratings may have indeed caused timbre to take a higher level of perceptual importance as musician's analytical listening mode was not necessarily engaged by the task.

Many of these uncertainties may also be a result of several limitations in the experimental work to date, most notably, the lack of diversity of stimuli. Most studies limit their study to two pitch-intervals (at most), with two timbres (usually piano and trumpet). These limitations leave many questions unanswered, particularly in terms of what timbral features in fact interfere with pitch. The prevailing theory that spectral changes are responsible for timbre's interaction with pitch has never been completely verified because no other study has investigated possible interactions with pitch and other dimensions of timbre, such as attack time and decay.

Resulting Goals and Hypotheses for Experiment

The current study therefore aims to address several of these issues. The primary goal is to better understand the role of musical training in pitch-timbre interactions. For this reason, a musician population with quantified interval-identification abilities was used as subjects in order to better control for training and ability. Musicians possessing relative pitch were used to ensure that the same type of categorical labelling was applied to interval identification, and also because the differences between relative pitch and perfect pitch possessors in pitch-timbre interactions is not fully understood (Marvin & Brinkman, 2000). This also allowed for control of the type of task

used to measure interactions. Many of the studies use a wide variety of classification and rating tasks, making it difficult to understand the nature of the differing results between studies.

The second goal was to increase the number of stimuli employed by using all intervals from the minor second to the octave in both ascending and descending directions and several instrumental timbres. By increasing the variety of timbres used in the experiment, it was possible to investigate how change along different dimensions of timbre (such as attack time and fluctuations in the temporal envelope) might interact with pitch. These factors have not been investigated by other studies.

Based on the previous research, three hypotheses are proposed:

- (1) Change along the dimension of spectral centroid will cause the most interference, as opposed to other timbral dimensions such as attack time, because spectral centroid covaries with pitch.
- (2) Pitch intervals that are identified the least accurately in a baseline condition will be more prone to interactions with timbre. This hypothesis is supported by the research showing differences between musicians and non-musicians (Beal, 1985; Pitt, 1994, Russo & Thompson, 2005b). If pitch processing is indeed the deciding factor on how much timbre is able to interfere with pitch, then the intervals that are most poorly identified should be more susceptible to interference with timbre.
- (3) Interactions with timbre will be revealed, particularly interval illusion (Russo & Thompson, 2005b), by consistent miscategorizations of intervals in a specific response category. For example, a tritone that has an increase in spectral centroid from the first to the second pitch may be labelled as a perfect fifth more often than a timbre-neutral tritone.

Pitch-Interval Perception

In order to better understand categorical interval perception and identify which intervals might be more susceptible to interference with timbre, it is necessary to examine the literature on interval identification. Not many studies have investigated categorical interval perception in musician populations. The few that have employ several different techniques to understand how intervals are identified.

Russo & Thompson (2005a) aimed to investigate differences between musicians and non-musicians in estimating the size of pitch intervals for intervals ranging from half a semitone to two octaves. Subjects were asked to rate the intervals on a scale from 1 (half a semitone) to 100 (two octaves) in both ascending and descending directions. Overall, there were effects of interval direction and register, indicating that descending intervals were rated as larger than ascending in lower registers, and ascending intervals were rated larger than descending intervals in the upper register. Musicians also demonstrated finer discrimination of intervals within the octave compared to non-musicians, although their ratings were similar to non-musicians for intervals over an octave. This result could indicate that musicians are more like non-musicians in their perception of intervals greater than an octave, and that subjective ratings of interval sizes can be influenced by the register in which intervals are presented. This finding suggests that intervals over the octave could be more susceptible to interference with timbre, and that register might affect the tendency to label intervals as larger than they are depending on the direction of the pitch change.

In his 1985 study, Andrzej Rakowski used a tuning exercise to examine how musicians tuned intervals made out of pure tones within the octave. The subjects were asked to tune a

variable tone against another fixed tone in order to produce the proper tuning of the interval given. Contrary to Russo & Thompson (2005a), results showed no effect of register on ratings, but found an effect of interval. Generally, there was a tendency to reduce the size of smaller intervals, and increase the size of larger intervals compared to their equal-tempered targets. Rakowski also found that major intervals (e.g., major sixth) were tuned as larger compared to their equal-tempered targets than their minor counterparts (e.g., minor sixth). Rakowski suggests that musicians tuned intervals based on musical context, particularly following their traditional harmonic resolution patterns. This result demonstrates that even when musicians are not required to use their explicit knowledge of tonal musical structures, their training influences pitch-interval judgments regardless of the task. It also suggests that timbre might interfere differently with major intervals (possibly labelled as larger) compared to minor intervals (possibly labelled as smaller).

Only a few studies have investigated categorical interval identification in musician populations. Killam, Lorton, & Schubert (1975) studied musicians' interval identification accuracy on melodic (ascending and descending) as well as harmonic intervals, from the minor second to octave. Response times were not collected. Their results demonstrate that harmonic (or simultaneous intervals) were the most difficult to identify, with ascending and descending melodic intervals being equally difficult. The most difficult intervals overall to identify were the minor sixth (55% accuracy), the minor seventh (58% accuracy), and major seventh (70% accuracy), followed by the tritone (at 72% accuracy). The most accurate intervals were the octave (88% accuracy) and major third (84% accuracy), followed by the minor second (83% accuracy). The perfect fourth, perfect fifth and major sixth were all equally discriminable at 82% accuracy, and the major second and minor third at 80% accuracy. Confusion matrices were also

computed for the most common errors. The data reveal that most intervals are likely to be confused with an interval a semitone away (on either side of the correct answer), and not with inversions or same class intervals (e.g., a major sixth was more likely to be labelled a minor seventh or minor sixth than a minor third). Exceptions to this were the ascending minor sixth, which was most often confused with the perfect fourth. The authors suggest a “Gestalt” effect of hearing a second inversion minor triad or first inversion major triad. The other confusion that was surprising was the descending perfect octave was most often confused with the perfect fifth. One other interesting finding was that there was little consistency among participants. Accuracy scores ranged from 50% to 95% accuracy overall. This was curious as all participants were undergrads in the music program who had completed computer-assisted instruction in interval recognition. Subjects who had higher accuracy (95% overall) were more likely to perform well on minor sixths (100%) than those who had lower scores. A subject-by-interval interaction was shown to be significant in all cases. This wide range of accuracy scores, as well as the differences in correct vs. incorrect intervals for this group of subjects suggests that even within musician populations, interval identification can vary greatly, making the need for quantifying pitch-interval identification very important for any studies that investigate interactions between pitch and timbre.

In another study, Art Samplaski (2005) specifically looked at interval confusions in melodic (ascending and descending) and harmonic intervals within the octave. Similarly to the Killam *et al.* (1975) study, they presented subjects with intervals in melodic or harmonic formation from the minor second up to the major seventh. These sounds were presented on a pseudo-clarinet timbre and not piano. Each interval was presented at ten different pitch levels at pitches from G4 to F5. The results replicates Killam *et al.* (1975) in that larger intervals were the

most difficult, with the minor sixth as most difficult. Confusions were most likely to occur between diatonic variants a semitone apart (e.g., minor second and major second, rather than major second with minor third). In contrast to Killam *et al.* (1975) who found that intervals were likely to be confused with those a semitone on either side (up or down), Samplanski found that larger intervals were more likely to be confused with intervals below them (e.g., the minor seventh confused with the major sixth and minor sixth more often than with the major seventh). These findings are contrary to the results found by Rakowski (1985), which showed in a tuning exercise that musicians increased the size of larger intervals and decreased the size of smaller intervals. These discrepancies demonstrate how important the type of task is for interval judgments, and that the effect of task should be taken into account when generalizing results.⁵

These results suggest that timbre is more likely to interfere with larger intervals, particularly the minor sixth, minor seventh and major seventh, as they were shown to be the most difficult to identify in Killam *et al.* (1975) and Samplanski (2005). The mode of presentation should have little effect since in the previous studies, ascending and descending intervals were as easily identifiable (contradicting Russo & Thompson's, 2005b, explanation for the musician's asymmetrical interval illusion in the descending direction). The research suggests that interval illusion could be seen by an increase in interval confusions⁶, which will most likely occur a semitone on either side of the interval class, as these are the most common confusions shown in timbre-neutral interval identification (Killam *et al.*, 1975; Samplanski, 2005).

⁵ This applies to many of the studies addressed in the literature review. The methods of many of the studies differ so widely that it becomes difficult to form hypotheses surrounding the phenomenon of pitch-timbre interactions. I have attempted to formulate hypotheses based on the literature available, but it should be noted that differences in method and task can only provide some insight into predictions and explanations for the experiment completed in this study, which differs quite drastically in task and data collection compared to previous studies.

⁶ Interval miscategorizations, e.g. minor sixths consistently being labeled as major sixths on certain trials. This could indicate, given the direction and timbre change that the increase in mislabeling could stem from an illusion effect.

2 Experimental Investigation

2.1 Aims and Hypotheses of Current Study

The current study aimed to investigate pitch-timbre interactions in a musician population⁷ with quantified pitch processing abilities. A timbre-neutral baseline procedure for interval identification was used to provide both accuracy and response time data for melodic intervals that are also used in trials in which the timbre changes between notes. I also wanted to investigate whether or not pitch-timbre interactions occurred when participants were asked to explicitly identify intervals. As many previous studies used subjective rating and/or comparison tasks, it is possible that musicians' categorical knowledge of melodic intervals played little importance in their responses, and thus allowed for timbre to play a more important role in subjective and comparative ratings. As a result, the goal was to investigate whether or not pitch-timbre interactions are robust enough to interfere with explicit, categorical identification of pitch-intervals, using traditional labels learned by musicians (e.g., perfect fifth, minor second, etc.). Lastly, I wanted to increase the amount of stimuli compared to what has been used in previous studies, which have primarily used only two timbres and two interval types. I wanted to test all melodic intervals⁸ within the octave, in ascending and descending directions, and include more acoustic timbres, with the goal of investigating whether changes along other dimensions of timbre (such as attack time or decay) interact with pitch in a similar manner to spectral centroid. Using recordings of acoustic timbres would also allow me to more easily apply the results of the study to music-theoretical analyses given their increased ecological validity compared to synthesized stimuli.

⁷ See participant information pages 22 and 27.

⁸ Sequential pitches, not simultaneous.

It was hypothesized that (1) change along the dimension of spectral centroid would interfere with pitch more than change along other timbral dimensions, such as attack time, because spectral centroid strongly covaries with pitch. Because musicians have been shown to be more susceptible to timbral changes when their pitch processing abilities are not able to aid them (Russo & Thompson 2005b, Beal 1985), it was also hypothesized that (2) intervals identified more poorly in the timbre-neutral baseline (having lower accuracy and/or slower response times) would be more susceptible to interference with timbre, demonstrated by even lower accuracy and/or slower response times in timbre-changing trials. And lastly, that (3) interval illusion (as discussed previously in Russo & Thompson 2005b) could occur in trials including timbre-change along the spectral dimension, which could be observed as a decrease in accuracy within timbre-change trials, coupled with an increase in miscategorizations in a specific response category and direction (e.g., an increase in tritones being mislabelled as perfect fifths in spectral trials with congruent timbre change).

2.2 Experiment 1: Timbre Selection on the Basis of Dissimilarity

A. Rationale

One of the primary objectives for this project was to investigate whether other dimensions of timbre (such as attack time) provided interference effects on interval identification. Changes in spectral centroid have been commonly referenced as the cause of pitch-timbre interactions (Melara & Marks, 1990a,b,c; Krumhansl & Iverson, 1990, Pitt 1994; Russo & Thompson, 2005b), while no other studies have investigated other dimensions of timbre. In order to ensure that the sounds varied along perceptually relevant timbral dimensions, a dissimilarity experiment

was completed on 16 orchestral sounds in order to complete Multi-Dimensional Scaling⁹ using CLASCAL (Winsberg & De Sote, 1993), with the extraction and verification of timbral descriptors using the TimbreToolbox in MATLAB (Peeters, Giordano, Susini, Misdariis, & McAdams, 2011). This experiment ensured that the timbre-pairs chosen would vary greatly along one timbral dimension, while varying as little as possible along other timbral dimensions.

B. Methods

Participants

Participants were recruited from the McGill University campus using an email notification. Nineteen participants (7 males and 12 females) took part in the study, ranging from 18 to 47 years in age. The mean age was 25 years ($SD = 6.5$). There were four non-musicians, only one of which had played an instrument before (4 years total on drum kit). There were fifteen musicians with a mean of 11.87 years of musical training ($SD=3.60$), who also had a mean of 3.93 years of ear training ($SD=3.04$) and 4.53 years of harmony training ($SD=2.53$). The primary instruments reported by the musicians were piano, saxophone, guitar, double bass, violin, viola, trumpet, voice, bass clarinet, bassoon and tuba. All participants completed a hearing test and had normal hearing at the time of the test. All participants read and signed a consent form before participation. The experiment conformed to the certification for ethic compliance under McGill Review Ethics Board (Certificate 67-0905).

Stimuli

The stimuli were selected from the Vienna Symphonic Library (VSL), and included flute, B-flat clarinet, oboe, English horn, bass clarinet, bassoon, French horn, tenor trombone, trumpet, muted

⁹ A method for visualizing similarities and differences between objects, represented visually as objects positioned in a space with a certain number of dimensions. Distance in the space models the perceived dissimilarities.

trumpet, violin, cello, marimba, harp, and vibraphone, as well as piano from August Förster samples. The sounds selected from the VSL were all mezzo forte long tones without vibrato, except for the tenor trombone which was selected at mezzo piano. The August Förster piano sounds were also chosen at mezzo piano. Both of the string sounds were bowed, and the marimba used hard mallets, while the vibraphone used extra soft mallets. All sounds were shortened to 700 ms using a 50ms raised cosine cut off, leaving the attack portion of all sounds intact. For each instrument, the pitch F-sharp 4 was used because it was the pitch directly in the middle of the register selected for the full experiment. Before the dissimilarity experiment, all sounds were equalized in loudness with respect to a comparison sound (oboe F-sharp 4) by ten people listening over Sennheiser HD280 Pro headphones. Median decibel values were taken from loudness matching and applied to the set of sounds. Sound levels were measured with a Bruel & Kjaer Type 4153 artificial ear to which the headphones were coupled, placed at the level of the listener's ears (Bruel & Kjaer, Nærum, Denmark). The stimuli were in the range of 50-56 db SPL. These sounds were then used in a short dissimilarity experiment where participants ranked the sounds in pairs as “identical” or “very dissimilar” on an unmarked scale that recorded dissimilarity as values between 0 (identical) and 1 (very dissimilar).

Apparatus

Sounds were stored on a Mac Pro 5 computer running OS 10.6.8 (Apple Computer, Inc., Cupertino, CA) and were amplified through a Grace Design m904 monitor (Grace Digital Audio, San Diego, CA) and presented over Sennheiser HD280 Pro earphones (Sennheiser Electronic GmbH, Wedemark, Germany). The experimental session was run with the PsiExp computer environment (Smith, 1995). Listeners were seated in an IAC model 120act-3 double-walled audiometric booth (IAC Acoustics, Bronx, NY).

Procedure

Participants were asked to rate the pairs of sounds on a scale from identical to very dissimilar. All 16 stimuli were presented in pairs and resulted in a total of 136 trials. There were two 68-trial blocks, allowing for a 5 minute break in between. Participants were able to hear all stimuli once in a randomized order once at the beginning of the first block. A short six-trial practice session was also conducted before the experiment began, using VSL sounds not included in the experimental trials (tuba, viola and string bass). The experiment took approximately 30 minutes, and participants were compensated \$10 for their time. Before the experiment, participants passed a pure-tone audiometric test at octave-spaced frequencies from 125 Hz to 8 kHz (ISO 389–8, 2004; Martin & Champlin, 2000) and were required to have thresholds at or below 20 dB HL to proceed to the experiment.

C. Results

A multidimensional scaling analysis using CLASCAL was completed on all dissimilarity ratings to obtain a three dimensional representation of the sounds used (Winsberg & De Soete, 1993; McAdams, Winsberg, Donnadieu, De Soete & Krimphoff, 1995, see also Figure 2.1). The analysis was completed in two stages: firstly, a clustering analysis of participant's dissimilarity ratings was completed in order to separate participants into groups that used similar rating strategies, and to eliminate any participants that did not use a systematic rating strategy (McAdams *et al.*, 1995). In the second stage, an analysis was completed using CLASCAL determined the number of latent classes, and the number of perceptual dimensions, and their weighting for each latent class (McAdams *et al.* 1995). Using the Timbre Toolbox in MATLAB, acoustic correlates were found for the resulting MDS dimensions found by the CLASCAL analysis (Peeters, *et al.*, 2011, see also Table 2.1). Correlations for all of different timbral

descriptors were computed for each pair of timbres. From these, those with the highest correlational values were selected. The x-axis was shown to correlate highly with temporal features, specifically Effective Duration (representing the perceived duration of the signal) and Temporal Centroid (the temporal centre of gravity of the energy envelope). The y-axis was shown to correlate highly with Spectral parameters, such as Spectral Skew and Spectral Centroid (spectral centre of gravity). The z-axis correlated most highly with Amplitude Modulation (modulation of energy over time). See Table 2.1 for a list of the most highly correlated descriptors.

From this space, three timbre pairs were chosen, which varied the most along one dimension, while varying the least along the other two dimensions (see Table 2.2). These pairs were French Horn-Piano for the temporal dimension, Marimba-Violin for the spectral dimension, and Bass Clarinet-Muted Trumpet for the other temporal dimension (amplitude modulation). The values shown in each column are the differences of the two timbres along the respective dimension (i.e., the space between the two sounds within the 3-dimensional space).

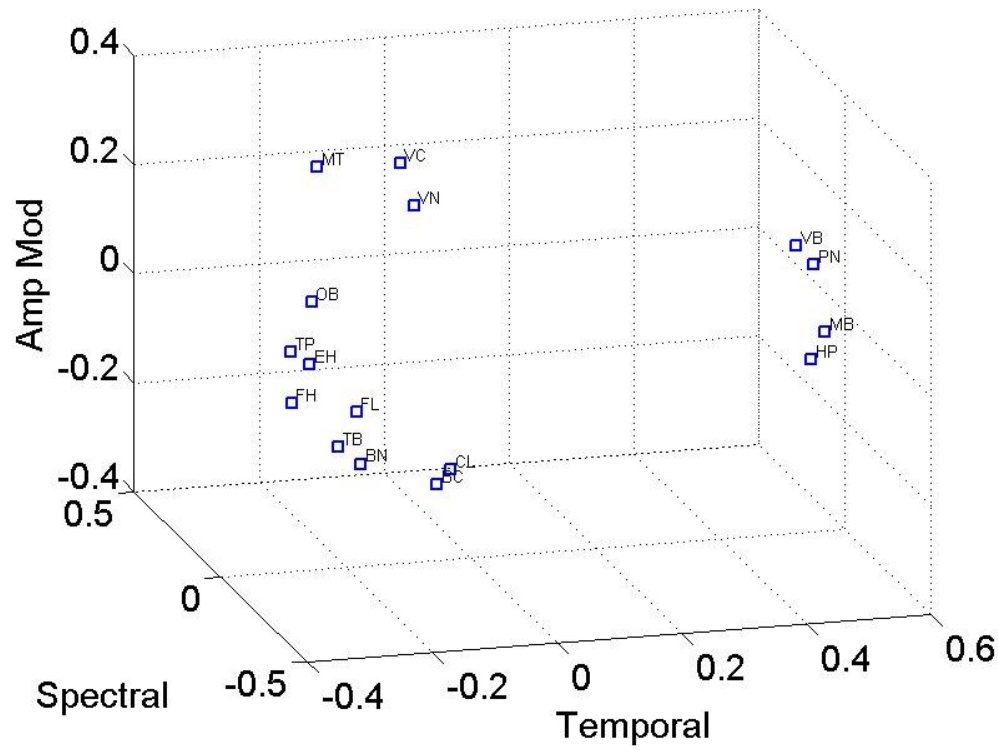


Fig. 2.1. 3D Timbre Space (FL=flute, CL=Clarinet, BC=Bass Clarinet, TB= Trombone, FH=French Horn, EH=English Horn, TP=Trumpet, OB=Oboe, MT=Muted Trumpet, VN=Violin, VC=Cello, VB=Vibes, PN=Piano, MB=Marimba, HP=Harp)

Table 2.1: Timbre Toolbox Descriptors and Correlation Coefficients

X-Temporal		Y-Spectral		Z-Temporal	
Decay Slope	-0.834	Frame Energy (median)	0.878	Amplitude Modulation	0.76
Temporal Centroid	-0.847	Spectral Centroid (median)	0.754		
Effective Duration	-0.943	Spectral Skew (median)	-0.755		
ERB Frame Energy (median)	-0.756	Spectral Slope (median)	0.754		
		Spectral Crest	-0.76		
		Decay Slope	0.753		

Table 2.2: Timbre Pairs and 3D Space Coordinates Differences

Timbre Pair	X Difference	Y Difference	Z Difference
FH-PN	-0.804188109	0.11632333	-0.220716332
VN-MB	-0.427718239	0.831993931	0.009797934
BC-MT	0.165782623	-0.094396772	-0.568413024

2.3 Experiment 2: Interval Identification

A. Rationale

The interval-identification experiment was constructed to record both accuracy and response times for interval identification of melodic intervals within the octave. Neither of the studies that have investigated interval identification collected response times (Killam *et al.*, 1975; Samplanski, 2005). Meanwhile, response times have been shown to be a key factor in demonstrating pitch-timbre interactions within musician populations because in many studies, musicians have reached ceiling effects for accuracy (Krumhansl & Iverson, 1992; Pitt, 1994). I therefore wanted to record accuracy and response times for interval identification as it was unclear if accuracy results alone would demonstrate interference with pitch, especially in highly trained musician populations.

B. Method

Participants

Participants were all recruited from McGill University, the University of Montreal, and the Université du Québec à Montreal. Most of the participants were from the McGill Schulich School of Music, and included a mixture of undergraduate and graduate students, all with relative pitch (self-screened). A total of 22 students, $M = 22$ years of age, $SD = 3.8$, 11 males and 11 females, took part in the experiment. All participants were trained musicians and had been

playing instruments for 8-21 years, $M = 15$ years, $SD = 3.4$. All musicians were currently active as performers, although playing time per week varied greatly in the range of 1-40 hours/week. Most musicians played more than one instrument. Musicians self-reported that their primary instruments were piano ($n = 7$), guitar ($n = 2$), violin ($n = 2$), oboe ($n = 1$), bassoon ($n = 1$), trumpet ($n = 2$), percussion ($n = 1$), trombone ($n = 1$), clarinet ($n = 1$), double bass ($n = 1$), cello ($n = 1$), and voice ($n = 2$). Secondary instruments included piano ($n = 5$), voice ($n = 3$), saxophone ($n = 2$), violin ($n = 3$), guitar ($n = 2$), electric bass ($n = 3$), trumpet ($n = 1$), cello ($n = 1$), oboe ($n = 1$), trombone ($n = 1$), percussion ($n = 2$), and flute ($n = 1$). Participants had a mean of 7 years of aural skills training, $SD = 4.5$, a mean of 5 years of training in harmony, $SD = 3.0$, and a mean of 4 years of training in musical analysis, $SD = 2.4$. Eleven of the participants were undergraduates still enrolled in aural skills, with one student in first year, three students in second year, and five students enrolled in third year post-tonal aural skills. All students, including those undergraduate and graduate students who had finished aural skills, had all done interval practice as part of their music education ($M = 5$ years, $SD = 4.1$). Nine participants had used software for ear training (1-3 years), and five of these had specifically used software for interval identification practice. All but one participant was trained in solfège, 3 knew moveable *do*, 11 knew fixed *do*, and 9 had been trained on both moveable and fixed systems. All participants completed a hearing test at the time of the experiment, and all had normal hearing. Participants read and signed a consent form. The experiment conformed to the certification for ethical compliance under McGill Review Ethics Board (Certificate 67-0905).

Stimuli

All 13 intervals were used (unison through to the octave) within the span of the 17 semitones from A#3 to D5. Four of each interval were selected for each interval quality (unison, minor second, major second, etc.), spread equally over the full range of the 17 semitones, resulting in a total of 48 intervals (see Table 2.3). Each pitch in the 17-semitone range was equally represented in the selection of intervals (see Table 2.4). Due to an error, one of the minor thirds was replaced as a perfect fifth, resulting in a total of only three minor thirds and five perfect fifths. This did not affect the pitch distribution over the register chosen. Each interval was presented twice, once in the ascending and once in the descending direction (excluding the unison) resulting in a total of one hundred trials. The timbre pairs selected from the CLASCAL analysis were French Horn-Piano (FH-PN, temporal), Marimba-Violin (MB-VN, Spectral), Bass Clarinet-Muted Trumpet (BC-MT, temporal, amplitude modulation), as well as the baseline Piano-Piano (PN-PN, neutral). The 100 intervals were presented in each timbre condition, with timbre-changing pairs presented in both directions (e.g., FH-PN as well as PN-FH). This resulted in a total of 700 trials: 100 piano-only trials, and 600 trials with timbre-change. Loudness matching was completed again for all selected sounds for all intervals over speakers by ten people. Mean dB values across participants were taken and applied to the complete set of stimuli. Sound levels were measured with a Brüel & Kjær Type 2205 sound-level meter (A-weighting) placed at the level of the listener's ears (Brüel & Kjær, Nærum, Denmark). The stimuli were in the range of 72-84 db SPL.

Table 2.3: Interval Distributions

Intervals	Low				High
Unison	A#3-A#3	F4-F4	G#4-G#4	C5-C5	
minor second	A#3-B3	D4-D#4	G4-G#4	B4-C5	
Major Second	C#4-D#4	E4-F#4	A4-B4	C5-D5	
minor third	...	C#4-E4	E4-G4	G4-A#4	
major third	B3-D#4	D4-F#4	G4-B4	A#4-D5	
perfect fourth	A#3-D#4	C4-F4	E4-A4	G#4-C#5	
tritone	B3-F4	E4-A#4	F#4-C5	G#4-D5	
perfect fifth	A#3-F4 B3-F#4**	C4-G4	D4-A4	F4-C5	
minor sixth	B3-G4	C#4-A4	D#4-B4	F#4-D5	
major sixth	B3-G#4	C4-A4	D4-B4	E4-C#5	
minor seventh	B3-A4	C4-A#4	C#4-B4	D#4-C#5	
major seventh	A#3-A4	C4-B4	D4-C#5	D#4-D5	
perfect octave	A#3-A#4	C4-C5	C#4-C#5	D4-D5	

Table 2.4: Pitch Distributions

Pitch	# of Occurrences
A#3	6
B3	7
C4	6
C#4	5
D4	**5
D#4	6
E4	6
F4	6
F#4	**7
G4	6
G#4	6
A4	7
A#4	6
B4	7

Apparatus

Sounds stored on a Mac Pro 5 computer running OS 10.6.8 (Apple Computer, Inc., Cupertino, CA) were amplified through a Grace Design m904 monitor (Grace Digital Audio, San Diego, CA) and presented over Dynaudio BM6a loudspeakers (Dynaudio International GmbH, Rosengarten, Germany) arranged at $\pm 45^\circ$, facing the listener at a distance of 1.5 m. The experimental session was run with the PsiExp computer environment (Smith, 1995). Listeners were seated in an IAC model 120act-3 double-walled audiometric booth (IAC Acoustics, Bronx, NY).

Procedure

The 700 trials were separated into 100 trial blocks, resulting in seven full blocks. The first block contained the baseline trials where only piano sounds were used. The following six blocks contained the timbre-change trials. All blocks were randomized in interval presentation order, with the final six blocks randomized in timbre-pair order as well. Each trial was 6 seconds,

with 1.55 seconds occupying stimulus presentation including a 50-ms inter-stimulus interval, and 4.5 seconds for participants to respond. Trials would time out after this time period and the trial would end. To move on to the next trial, participants would press the space bar on a computer keyboard. Participants were instructed to say the interval name out loud. Responses were collected using a Behringer ECM 8000 microphone. The responses for each trial were saved as a .wav file, which included the stimulus presentation and participant response. Participants were instructed to use any interval name with which they were comfortable, with the stipulation that the interval name must have included the interval size and quality. They were also allowed to speak in either French or English and were instructed to only give one answer per trial and to use the same interval labels throughout the experiment. Participants were told to respond as quickly as possible, but were urged not to answer until they were certain of their answer to avoid stuttering and unclear vocal responses. Participants were told to speak clearly and at a reasonable pace. No singing of intervals was allowed, and participants were highly discouraged from making any sounds outside of naming the intervals. A practice session of twenty-six sounds completed before the experimental blocks was also used as a pre-screening procedure. This practice session used a subset of trials from the experimental block on piano only, and included two of each interval, ascending and descending (with three perfect fifths and one minor third, due to the error mentioned above). In order to proceed to the full experiment, participants needed a minimum score of 18/26 correct on this practice session. Participants with a lower score were rejected from the full experiment, and compensated \$5.00 for their time. This screening procedure was implemented to ensure that participants spoke clearly enough for proper data collection, and also so that enough response time data could be collected for analysis, as response times would be analyzed for correct trials only. Before the experiment, participants

passed a pure-tone audiometric test at octave-spaced frequencies from 125 Hz to 8 kHz (ISO 389-8, 2004; iss) and were required to have thresholds at or below 20 dB HL to proceed to the experiment.

Evaluation of Trials

Each response was recorded as a .wav file and was evaluated using a custom-made evaluation program designed by the McGill MPCL Technical Manager Bennett K. Smith. The evaluation program would automatically detect the stimulus presentation start and ending, as well as the start and ending of speech. The screen shot in Figure 2.2 shows the interface of the evaluation program. The .wav file is shown at the top of the interface, which includes the stimulus presentation at the beginning, with the participant's spoken response following it. Interval labels can be seen on the far right hand side of the interface, with participant data on the far left. Individual trials are in the second column in on the left, which contain the stimulus names (containing the instrument name, plus MIDI number representing pitch frequency, for example, PN_59-PN_65 seen at the very top of the column representing piano B3 going to piano F4). The second column on the right contains the block numbers (1-7), with a box for entering comments below, and playback settings at the very bottom of that column. To evaluate each trial, the experimenter would first listen to the interval name spoken by the subject and select the interval name for each trial by clicking on the appropriate label on the far right hand side. The evaluation program would automatically determine if the interval was correct by comparing the difference between MIDI pitch numbers of the stimulus files used and the number of the interval label selected, which was recorded in interval-class number (numbers between -12 to +12 depending on the direction of the interval; this can be seen just above the comment box in the

second column in on the right). To accurately determine the response time, the experimenter would adjust the markers at the start and ending of speech (shown in blue highlighted area over the wave form), which was measured from the beginning of the second tone to the beginning of the speech. The duration of speech was also recorded by moving the end of the blue highlighted area to the end of speaking time. The trial would not be recorded in a .csv file until the marker placements were manually adjusted by the experimenter to ensure that response time was accurately evaluated by the experimenter. The value for measuring the start of response time was automatically set for each trial to 750 ms at the beginning of the second stimulus in the final .csv file to ensure that all response times were as accurate as possible. The program was designed to run in real time while participants completed the experiment in order to allow fast and accurate evaluation of the pre-screening test.

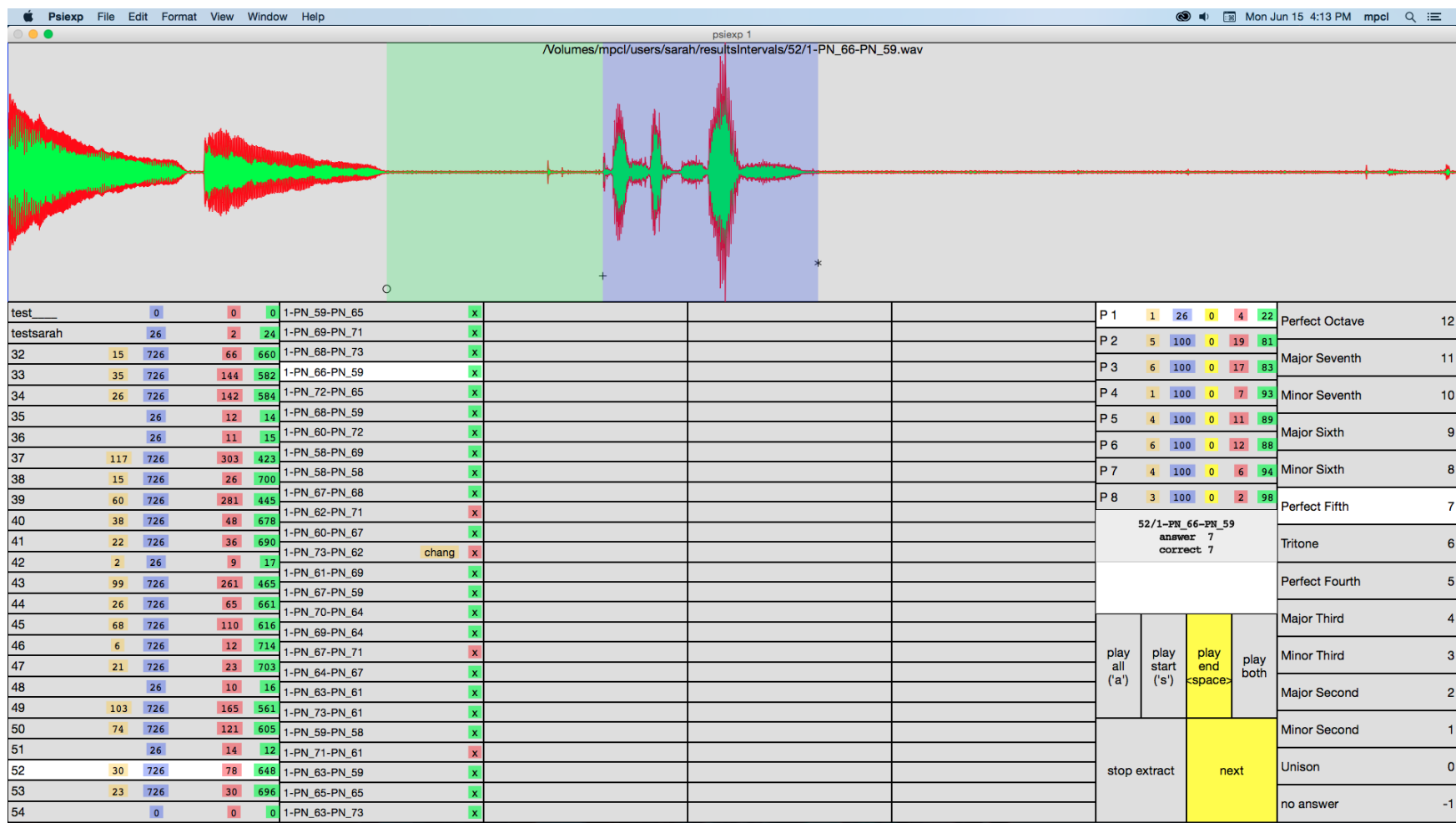


Fig. 2.2. Screen Shot of Evaluation Program

Speech Errors and Evaluation Method

Some common issues resulted from collecting speech data. These issues often fell into several types of responses, such as 1) multiple answers, 2) extended/elongated answers, 3) speech/sound before response, and 4) partially complete or timed-out answers.

The first category included responses where participants provided more than one answer, such as “Major third, no minor third.” In such an instance where two complete answers were given, the first answer was used as the response, with response time starting at the beginning of speech. Other issues in this category also occurred in the form of a change in quality such as “Major, minor third.” These responses often were blurred together such as “Maaaiinor third.” All responses where a change in quality occurred, either as two complete words, or one combined word, were marked as “no response,” marked as incorrect, and therefore not included in the response time analyses. A comment was added for these trials indicating what the change of interval was in case further analysis was needed.

The second category of response that proved difficult in evaluation included extended or elongated answers. These were responses such as “Mmmmmmmmajor third,” or “minor ssssseventh.” These trials were evaluated as normal trials, with response time recorded at the beginning of speech (whether the beginning was stretched out or not). A comment was placed on these trials indicating what part of speech was extended in order to tag them for future analysis. Because the length of speaking was also recorded, I anticipated that stretched trials would be longer than normal trials of the same intervals class. For this future analysis therefore, I plan on seeing if the number of stretched trials increases or decreases as a function of timbre-pair. For the third class of response issue, typically seen as extraneous sound before interval response (i.e., shifting in the seat, exhaling, etc.), the extraneous sounds were ignored, and response time was

measured from the start of the interval label response.

The final category of response issue included partially complete or timed-out answers. These were responses that were unfinished due to the trial timing out, such as “Major thir.” These responses were marked as correct if enough of the interval information was available to clearly identify what was said. In English and French, both “uni” and “octa” were acceptable answers for unison and octave. For the other intervals in English, enough of the interval number must have been said to accurately identify the spoken interval. This was particularly true for the sevenths, sixths and seconds (the participant must have responded with “sev,” “six” or “sec” for the interval to be included). In French, enough of the interval quality must have been available, such as “min” for minor, “maj” for major intervals, and “jus” (*juste*) for perfect intervals. If this information was not available, the trial was marked as “no answer” and was not included in the data analysis. A comment was placed on these trials indicating that the trial was partially timed out, and how much of the speech was included. For trials in which no answer was given, the trial was marked as “no answer,” with the word “none” was entered in the comment section.

Overall participants were able to easily respond vocally for interval recognition, with a small subset of trials presenting the issues mentioned above. If a trial proved too ambiguous to identify clearly, the trial was thrown out. Trials where nonsense answers were provided (such as “perfect sixth” or “major fifth”) were marked as “no-answer” with an appropriate comment identifying the error with the response. Participants were screened for speech clarity, as well as accuracy. No participant failed the pre-screening due to speech issues, only interval accuracy ended up being a factor for pre-screening.

C. Results

Accuracy

Firstly, a single-factor analysis of variance (ANOVA)¹⁰ with Timbre Pair as repeated measure was completed for the unison interval in order to see whether or not timbre pair had an effect on accuracy for unisons. There was no effect of timbre pair found, $F(6,126) = 1.12$, $p = .36$, $\eta_p^2 = .051$.¹¹ Accuracy was 97% and higher for unisons (see Table 2.5). For all other intervals, a repeated-measures ANOVA was completed using Timbre Pair (PN-PN, FH-PN, PN-FH, MB-VN, VN-MB, BC-MT and MT-BC)*Direction (ascending and descending)*Interval (12 total, minor second through to the octave) in order to observe general effects of each, as well as any systematic interactions between timbre, interval and the direction of the interval. No effect of timbre pair was found, $F(6,126) = 1.11$, $p = .36$, $\eta_p^2 = .050$ (see Table 2.6), indicating that timbre pair had no effect on participant accuracy for any intervals. Apart from timbre, there was a general effect of direction, $F(1,21) = 19.24$, $p < .001$, $\eta_p^2 = .478$ (see Table 2.7), indicating that descending intervals were generally less accurate than ascending ones. A general effect of interval was found, which was corrected for violations of sphericity using Greenhouse-Geisser epsilon, $F(4.08,85.75) = 11.41$, $\epsilon = .371$, $p < .001$, $\eta_p^2 = .352$ (see Table 2.8). This demonstrates that the least accurate intervals were the minor sixth and minor seventh, while the most accurate were consistently the octave, minor second, and minor third across all participants. No

¹⁰ Statistical test designed to analyze variance within and between groups (either independent groups for a single-factor ANOVA or related groups for a repeated-measures ANOVA).

¹¹ The symbol 'F' is the test statistic, while the numbers in parenthesis e.g., (1,2) next to it are the (1) degrees of freedom of the effect being tested, and (2) those of the error term, respectively. The symbol 'p' represents the probability that the difference being tested is zero. One generally considers $p < .05$ to be statistically significant. The epsilon symbol 'ε' (used later) is a measure of the departure from sphericity. Sphericity is the condition in which the variances of the differences between all combinations of related groups are equal. If $\epsilon < 1$, the degrees of freedom of the F-test are multiplied by this factor to make the test more conservative. The symbol ' η_p^2 ' or partial-eta-squared, is a measure of the size of the effect being measured.

interactions of Timbre Pair*Direction $F(1,126) = 1.14$, $p = .341$, $\eta_p^2 = .052$ or Timbre Pair*Interval $F(66,1386) = 1.24$, $p = .100$, $\eta_p^2 = .045$ were found, and no three-way interaction of Timbre Pair*Direction*Interval was found, $F(66, 1386) = .986$, $p = .511$, $\eta_p^2 = .045$, indicating no systematic interactions of timbre with interval and direction. Independent of timbre, an interaction of Direction*Interval was observed, $F(4.94, 103.80) = 2.55$, $\epsilon = .449$, $p = .032$, $\eta_p^2 = .108$ (see Fig. 2.3), showing that certain intervals, most notably the major sixth and minor seventh, were far more accurate in the ascending direction than they were in the descending direction. A subsequent paired t-test¹² was completed, comparing the means of each interval in both directions (i.e., compare ascending minor second with descending minor second), in order to see if these differences were statistically significant. The Bonferroni-Holm correction¹³ for multiple tests was also applied. Only one significant difference was found for the major seventh. The ascending major seventh ($M = .77$, $SD = .27$) was significantly more accurate than the descending major seventh ($M = .67$, $SD = .29$), $t(21) = 3.33$, $p = .003$.

Table 2.5: Mean Accuracy for Unisons by Timbre Pair

TimPair	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
PN-PN	1.000	0.000	1.000	1.000
FH-PN	1.000	0.000	1.000	1.000
PN-FH	1.000	0.000	1.000	1.000
MB-VN	.977	.023	.930	1.025
VN-MB	.977	.016	.945	1.010
BC-MT	1.000	0.000	1.000	1.000
MT-BC	1.000	0.000	1.000	1.000

¹² A statistical technique that is used to compare two population means in the case of two samples that are matched on all conditions.

¹³ Statistical test designed to reduce Type I errors (assumed statistical significance where none exists), which arise from multiple comparisons. The test increases the criterion needed for a result to be considered statistically significant.

Table 2.6: Mean Accuracy for Timbre Pair (all other intervals 1-12)

TimPair	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
PN-PN	.836	.032	.769	.903
FH-PN	.844	.028	.786	.902
PN-FH	.842	.031	.777	.906
MB-VN	.826	.035	.754	.898
VN-MB	.829	.035	.757	.901
MT-BC	.840	.033	.772	.908
BC-MT	.842	.031	.778	.905

Table 2.7: Mean Accuracy by Direction

Direction	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.863	.028	.805	.921
2	.811	.036	.737	.885

1: ascending 2: descending

Table 2.8: Mean Accuracy by Interval

Interval	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.942	.017	.907	.978
2	.884	.033	.815	.952
3	.927	.019	.888	.966
4	.864	.037	.788	.941
5	.875	.034	.804	.946
6	.860	.038	.781	.940
7	.888	.030	.825	.950
8	.719	.061	.592	.846
9	.778	.043	.687	.868
10	.659	.066	.523	.795
11	.719	.057	.600	.838
12	.927	.025	.876	.978

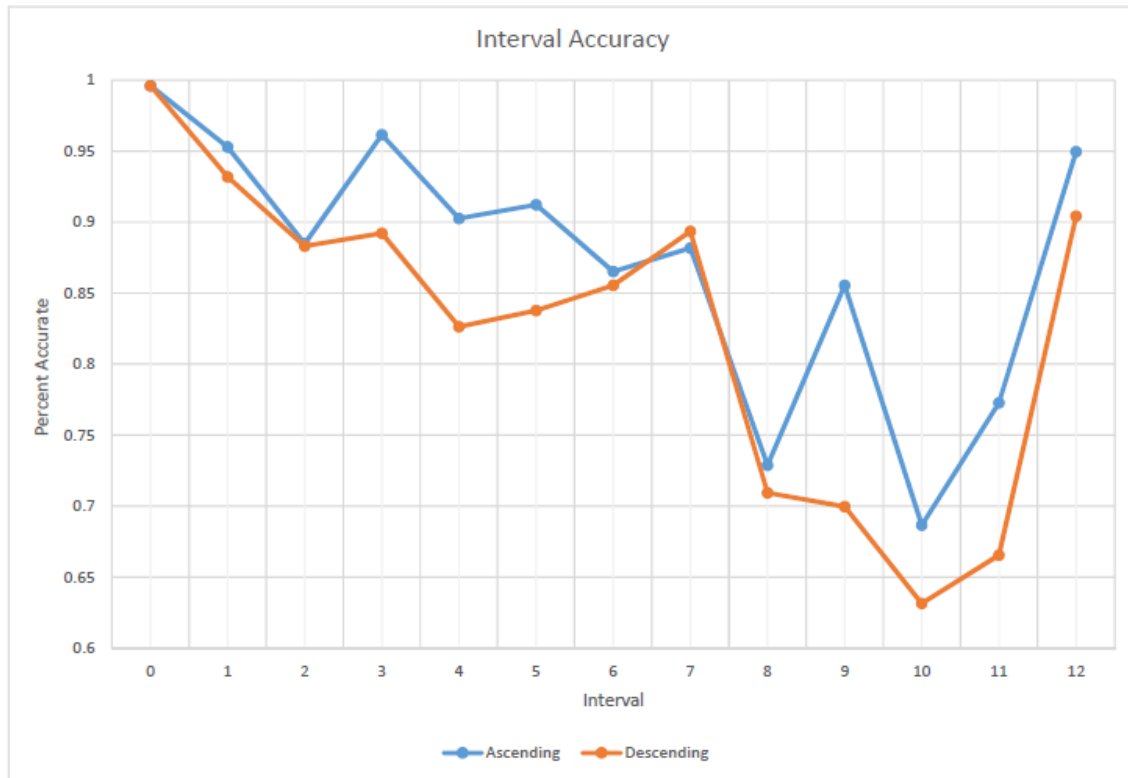


Fig. 2.3. Response Accuracy as a Function of Interval for the Two Interval Directions

A third repeated-measures ANOVA examined Order of Timbre Pair(2)*Direction(2)*Interval (12) separately for each timbre pair. This test compared, for example, the intervals of FH-PN to those of PN-FH to see if the ordering of the timbres affected the accuracy for any interval in either or both directions. Only one marginally significant ($.05 > p > .10$) interaction was found for the timbre pair FH-PN with Direction, $F(1,21) = 3.726$, $p = .067$, $\eta_p^2 = .151$ (see Fig. 2.4 below), demonstrating a slight improvement of accuracy when French Horn was the pitch on the bottom in either ascending or descending intervals across all intervals (was not specific to any particular interval class). No other significant effects of Order or its interaction with Interval were found.

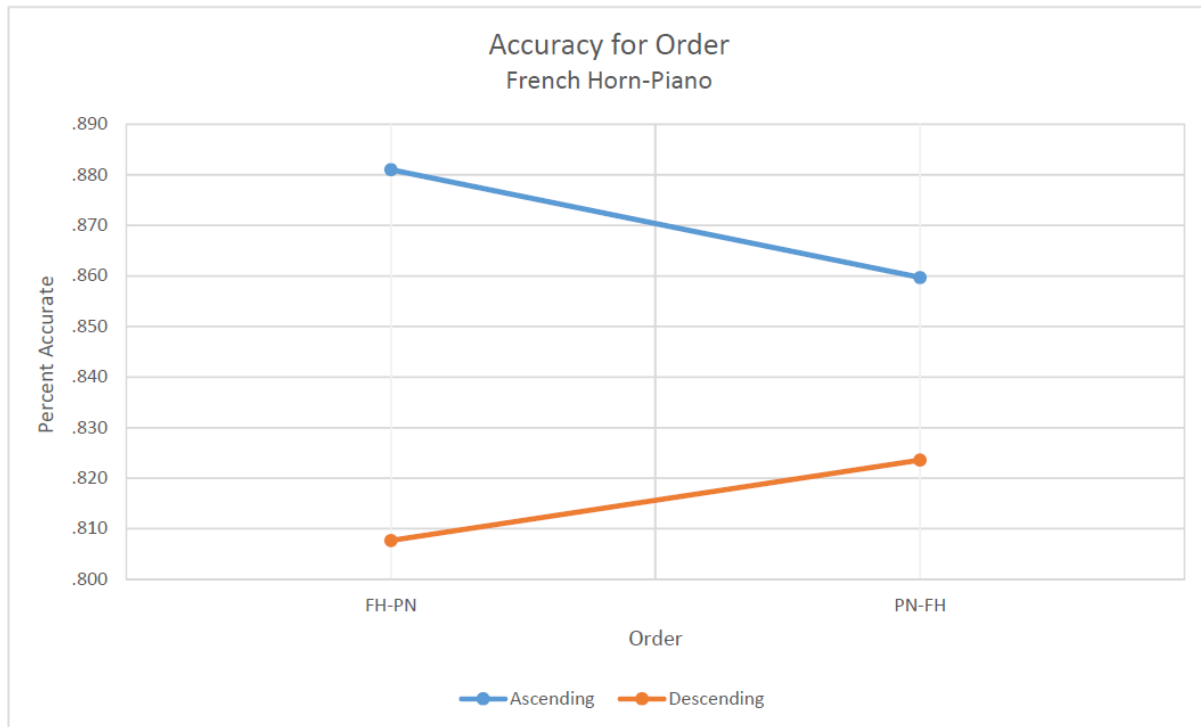


Fig. 2.4. Mean Accuracy (%) by order (FH-PN)

Although Timbre Pair was not shown to be significant for accuracy and did not interact with Interval or Direction, I wanted to investigate whether or not timbre pair affected the category of interval miscategorizations by participants (i.e., what interval labels participants used in incorrect trials). To check this, confusion matrices were computed for each timbre pair (see Appendices, Tables A.1-A.14). These tables should be read horizontally. The correct intervals are presented in rows, while the responses given by participants are presented in columns. Correct responses are therefore seen along the diagonal shown in green, while incorrect responses are highlighted in red. From this we can see, in the piano condition for example, that unison (in the first row), read horizontally, has no errors, while the minor second shown in the second row down, has miscategorizations as a major second and minor third. The values are given in proportions of trials shown between 0-1. Generally, the piano-only condition showed that most confusions cluster within one or two semitones of the correct answer, indicating that

when an error was made, it was most likely to be with an interval a semitone or tone away on either side. Exceptions to this are the minor second (which was never confused with a unison), the perfect fourth and perfect fifths (which were never confused with the tritone), the perfect fifth was also never confused with a minor sixth (only a perfect fourth), the tritone (which was most confused with the minor sixth, and never the major third, but rather with the minor third), the minor seventh (which was most likely to be confused with the minor sixth), the major seventh (which was never confused as an octave, and was more likely to be confused with the minor seventh and tritone), and finally the octave (which was more likely to be confused with the perfect fifth, major sixth and unison, and not the major seventh). Another general trend observed was that the range of confusions generally increased with the size of the intervals.

Miscategorizations were fairly limited in small intervals (with the exception of the major third), and were fairly large in the sixths, sevenths, and octave.

Overall, the confusion matrices for the timbre-changing pairs reveals less tight clustering of miscategorizations near the correct answers, and therefore an increase in the spread of miscategorizations, indicating that intervals overall were more likely to be miscategorized as intervals larger than a semitone or tone away. This is particularly true for the French Horn-Piano pairs and Marimba-Violin pairs. In both cases, intervals were more likely to be incorrectly labelled as more than a tone or semitone away. This can be seen in the FH-PN condition (see Appendices 2.3-2.4) with new miscategorization types that did not occur in the PN-PN trials (for example, the minor second being incorrectly labelled as a major seventh, the major third being incorrectly labelled as a major sixth, and the perfect fourth as an octave). The Marimba-Violin condition (see Appendices 2.7-2.8) also has some interesting and unique response errors. It is the only condition where unison errors were made, which were mislabeled as octaves (in the MB-

VN and VN-MB conditions) and once as a minor second in the VN-MB condition. A substantial increase in many intervals being incorrectly labelled as octaves was observed in both the MB-VN and particularly the VN-MB condition (all but the major second, minor third, tritone and major sixth were labelled as an octave at one point in either condition). The MB-VN condition also contains some of the only unison confusions, most notably the minor second, major second and major third are all incorrectly labelled as unisons in a portion of the error responses. These types of errors do not happen in any of the other timbre conditions, except for one unison confusion in the MT-BC condition where one major second was confused for a unison.

In order to verify these findings, root-mean-squared errors (RMSE)¹⁴ were calculated for each interval type in semitones for each timbre condition in order to assess the spread of errors off the diagonal (see Table 2.9 and Figure 2.5). The table shows that the overall mean values for each timbre pair differ quite a bit. The PN-PN condition has the smallest spread (mean RSME of 2.81), while the MB-VN pair has the largest spread (mean RSME of 4.29). The graph in Figure 2.6 shows that the RSME also varied by interval for each timbre pair, the most notable differences being the higher RSME for the unison for the MB-VN and VN-MB pairs only, as well as the high RSME for MB-VN in the minor sixth.

¹⁴ This statistic test is used to measure the average distance that a data point is from a fitted line. In this instance, it is the average distance of participant responses from the correct interval in number of semitones.

Table 2.9: Root Mean Square values for each timbre pair

Interval	Root Mean Square values by Timbre Pair (semitone distance from correct response)						
	PNPN	FHPN	PNFH	MBVN	VNMB	BCMT	MTBC
0	0	0	0	4.71	3.34	0.00	0.00
1	1.04	2.79	2.63	4.10	5.05	3.26	1.14
2	1.18	1.75	4.42	4.70	3.96	2.25	1.90
3	1.07	1.90	1.52	1.39	2.22	1.73	1.41
4	2.20	3.59	3.96	4.83	4.14	1.98	2.56
5	2.17	3.75	3.03	2.76	4.20	1.84	2.80
6	2.29	4.12	5.31	4.94	5.12	3.91	4.76
7	3.65	3.86	3.92	3.13	4.10	3.52	3.78
8	3.93	3.52	4.01	5.60	4.02	3.89	3.56
9	3.27	3.23	2.86	2.67	3.51	4.35	2.86
10	5.07	4.95	3.61	5.04	5.33	5.01	4.45
11	4.71	6.56	5.99	5.60	5.92	6.73	6.87
12	5.97	5.71	4.55	6.33	4.40	4.20	5.80
Mean RMS	2.81	3.52	3.52	4.29	4.25	3.28	3.22

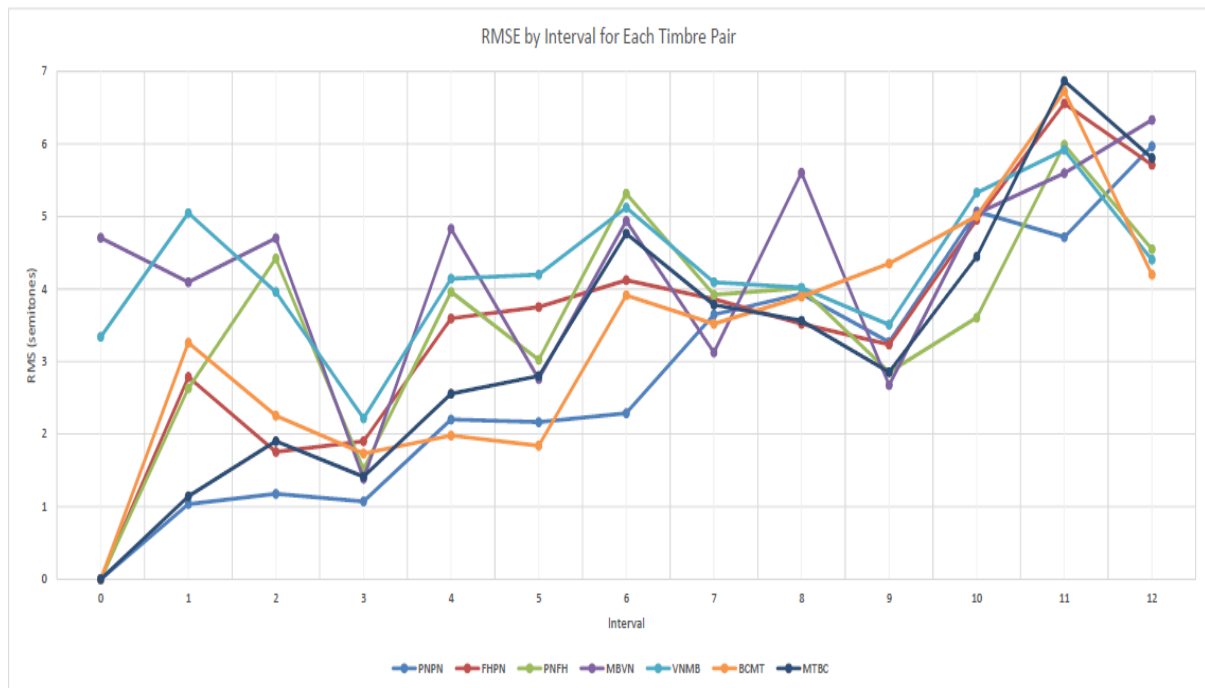


Fig. 2.5. Root-mean-squared error (in semitones) for each timbre pair.

Response Time

Eleven out of the 22 participants ended up having missing response time data due to incorrect and/or timed out trials. Even if participants were missing a response time value for one interval in one timbre condition (e.g., ascending major second in FH-PN), that participant would be excluded from the analysis because for the repeated-measures ANOVA, SPSS does list-wise deletion for any missing data. As a result, these missing response time data were replaced with the mean value (of log response times) over all other participants without missing values for a given interval class, direction and timbre pair (e.g., ascending minor second for MT-BC). Log response times were used because raw response times (in seconds) taken from perceptual experiments have been shown to be extremely skewed (Baayen & Milin, 2010). There were a total of 105 of missing response times, which represents less than 1% (0.68%) of the entire data set (105/15,400 trials). There were no more than four missing response times for any given interval category for any particular timbre pair. After this modification, the same ANOVAs were performed on the log response time data that were applied to the accuracy data. The ANOVA on unisons revealed no effect of timbre, $F(4.09, 85.85) = .929$, $\epsilon = .681$, $p = .453$, $\eta_p^2 = .042$ (see Appendix, Table B.1), indicating that timbre pair did not affect the response times for unisons. For the rest of the intervals, the ANOVA on Timbre Pair(7)*Direction(2)*Interval(12) revealed a significant main effect of Timbre Pair (see Table 2.10), $F(3.44, 72.16) = 5.63$, $\epsilon = .573$, $p = .001$, $\eta_p^2 = .211$. Table 2.10 shows the means of the log response times (left), with the response times in seconds (right). From this, we can see that the fastest timbre pair overall was not PN-PN, but in fact FH-PN, while the slowest was MB-VN. Similarly to the accuracy results, there were also main effects of Direction, $F(1, 21) = 88.89$, $p < .001$, $\eta_p^2 = .809$, and Interval, $F(11, 231) = 15.59$, $p < .001$, $\eta_p^2 = .426$ (see Tables 2.11-2.12). The Timbre Pair*Interval interaction was significant,

$F(66, 1386) = 1.59, p = .002, \eta_p^2 = .070$ (see Figure 2.6), as was the Direction*Interval interaction, $F(11, 231) = 2.46, p = .006, \eta_p^2 = .105$ (see Figure 2.7). Paired t-tests were completed for each interval pair comparing ascending and descending directions. After applying the Bonferroni-Holm method, significant differences were found for several intervals, indicating that these intervals were significantly faster in the ascending direction than in the descending direction (see Table 2.13). These included the minor second, major second, minor third, major third, perfect fourth, major sixth and octave. The Timbre Pair*Direction interaction was not significant, $F(6, 126) = .925, p = .479, \eta_p^2 = .042$, nor was the three-way Timbre Pair*Direction*Interval interaction, $F(66, 1386) = 1.05, p = .364, \eta_p^2 = .048$, indicating no systematic interactions between pitch, timbre and direction.

Table 2.10: Response Times for Timbre Pair (all other intervals 1-12)

TimPair	Mean (LogRT)	Std. Error	95% Confidence Interval		Mean (seconds)
			Lower Bound	Upper Bound	
PN-PN	.631	.036	.557	.705	1.879
FH-PN	.599	.044	.507	.691	1.819
PN-FH	.627	.042	.541	.714	1.872
MB-VN	.668	.040	.584	.752	1.95
VN-MB	.605	.042	.517	.692	1.83
BC-MT	.631	.042	.543	.718	1.879
MT-BC	.645	.042	.558	.731	1.905

Table 2.11: Response Times for Direction

Direction	Mean (LogRT)	Std. Error	95% Confidence Interval		Mean (seconds)
			Lower Bound	Upper Bound	
1	.575	.039	.493	.656	1.777
2	.684	.042	.597	.771	1.981

Table 2.12: Response Times for each Interval

Interval	Mean (LogRT)	Std. Error	95% Confidence Interval		Mean (seconds)
			Lower Bound	Upper Bound	
1	.451	.043	.361	.540	1.569
2	.572	.055	.458	.685	1.771
3	.585	.060	.459	.710	1.794
4	.634	.052	.526	.741	1.884
5	.698	.058	.578	.818	2.01
6	.564	.070	.418	.710	1.75
7	.667	.056	.551	.783	1.948
8	.813	.037	.735	.891	2.254
9	.721	.046	.625	.818	2.057
10	.764	.048	.665	.864	2.147
11	.735	.041	.649	.821	2.085
12	.348	.045	.255	.441	1.416

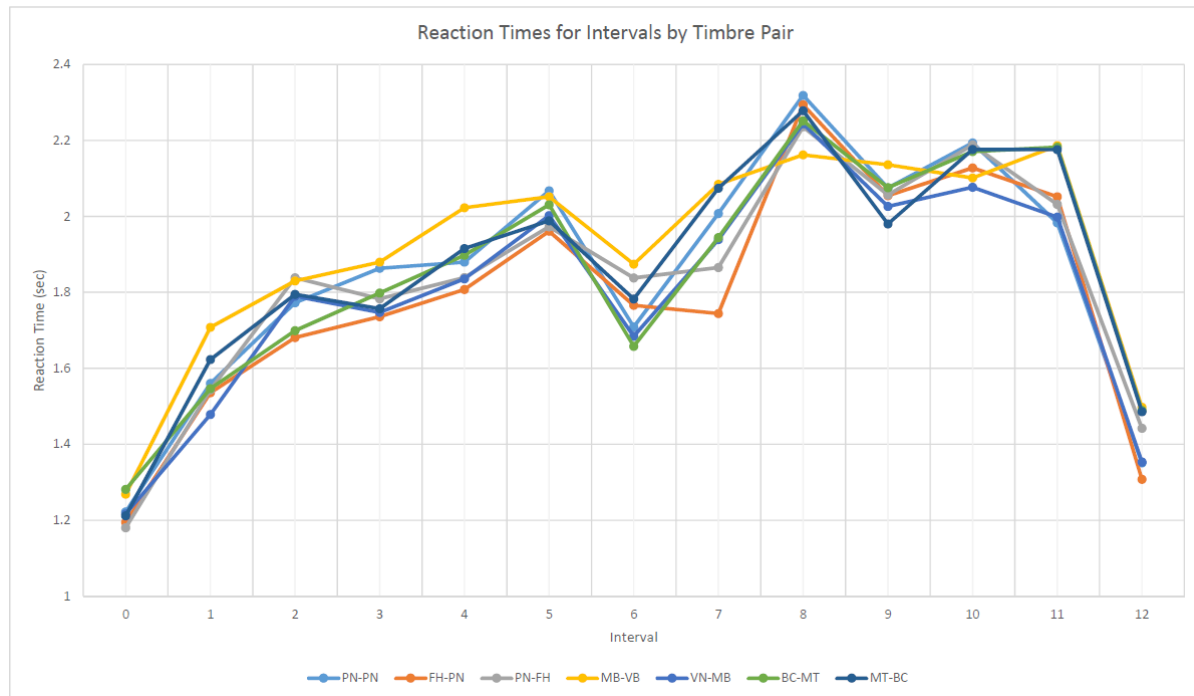


Fig. 2.6. Response times (in seconds) for the identification of different intervals for each Timbre Pair.

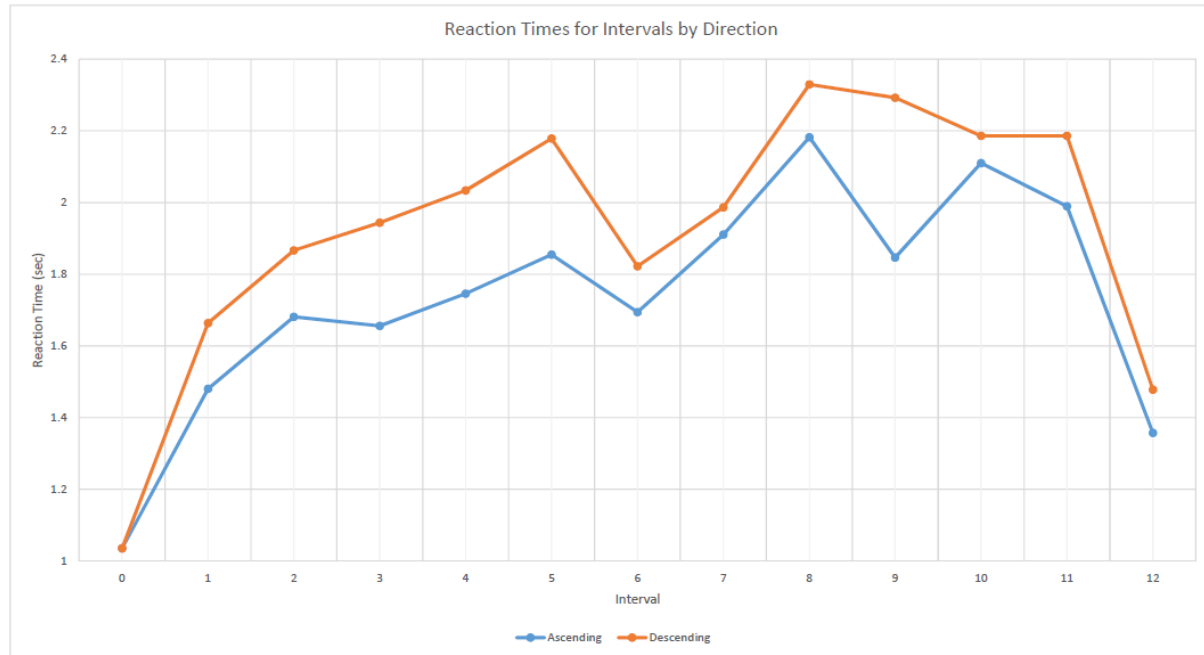


Fig. 2.7. Response times (in seconds) for the identification of different intervals for each direction.

Table 2.13: T-test results on response times for intervals by direction

Interval	Ascending	Descending	t-test
Minor Second	M = .392, SD = .197	M = .509, SD = .24	t(21) = -3.21, p = .004
Major Second	M = .519, SD = .258	M = .624, SD = .27	t(21) = -3.85, p = .001
Minor Third	M = .504, SD = .29	M = .665, SD = .308	t(21) = -3.81, p = .001
Major Third	M = .557, SD = .254	M = .71, SD = .273	t(21) = -3.46, p = .002
Perfect Fourth	M = .618, SD = .273	M = .779, SD = .284	t(21) = -5.59, p < .001
Major Sixth	M = .613, SD = .242	M = .83, SD = .216	t(21) = -6.8, p < .001
Octave	M = .305, SD = .209	M = .391, SD = .225	t(21) = -3.53, p = .002

Finally, Order(2)*Direction(2)*Interval(12) ANOVAs were performed on each timbre pair separately to see if the order of the timbres affected response times for intervals in either direction. Effects of order were found for all timbre pairs, indicating that one order for each timbre pair was faster than the other. For the French Horn-Piano pair, the marginally significant effect of order, $F(1, 21) = 4.04$, $p = .057$, $\eta_p^2 = .161$ (see Table 2.14) showed that FH-PN was

faster than PN-FH. Similarly to the accuracy data, a marginally significant Order*Direction interaction was also found for French Horn-Piano, $F(1, 21) = 3.93$, $p = .061$, $\eta_p^2 = .158$ (see Figure 2.8), indicating that the response time was faster in the descending direction when French Horn was on the bottom. For the Marimba-Violin pair, an effect of Order was found, $F(1, 21) = 33.97$, $p < .001$, $\eta_p^2 = .618$ (see Table 2.15), indicating that VN-MB was far faster than MB-VN. An Order*Interval interaction was also observed for the Marimba-Violin pair, $F(11, 231) = 1.97$, $p = .033$, $\eta_p^2 = .086$ (see Figure 2.9). Paired t-tests were completed for each interval comparing the two timbre presentation orders for the MB-VN pair (i.e., MB-VN vs. VN-MB for each interval). After Bonferroni-Holm correction, only two significant differences were found: one for the minor second (MB-VN, $M = .535$, $SD = .216$; VN-MB, $M = .391$, $SD = .211$), $t(21) = 5.21$, $p < .001$; the second for the octave (MB-VN, $M = .404$, $SD = .242$; VN-MB, $M = .302$, $SD = .232$), $t(21) = 3.43$, $p = .003$, indicating that the VN-MB pair was far faster than the MB-VN pair for these two intervals in particular. For the pair Bass Clarinet-Muted Trumpet, an effect of Order was found, $F(1, 21) = 3.25$, $p = .086$, $\eta_p^2 = .134$ (see Table 2.16), indicating that BC-MT was generally faster than MT-BC. All order effects can be seen in Table 2.17 converted into seconds. No other significant effects were observed, and no three-way interactions of Order*Direction*Interval were observed for any timbre pair.

Table 2.14: Mean Response Times by Order for FH-PN (LogRT's)

Order	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
FH-PN	.599	.044	.507	.691
PN-FH	.627	.042	.541	.714

Table 2.15: Mean Response Times by Order for MB-VN

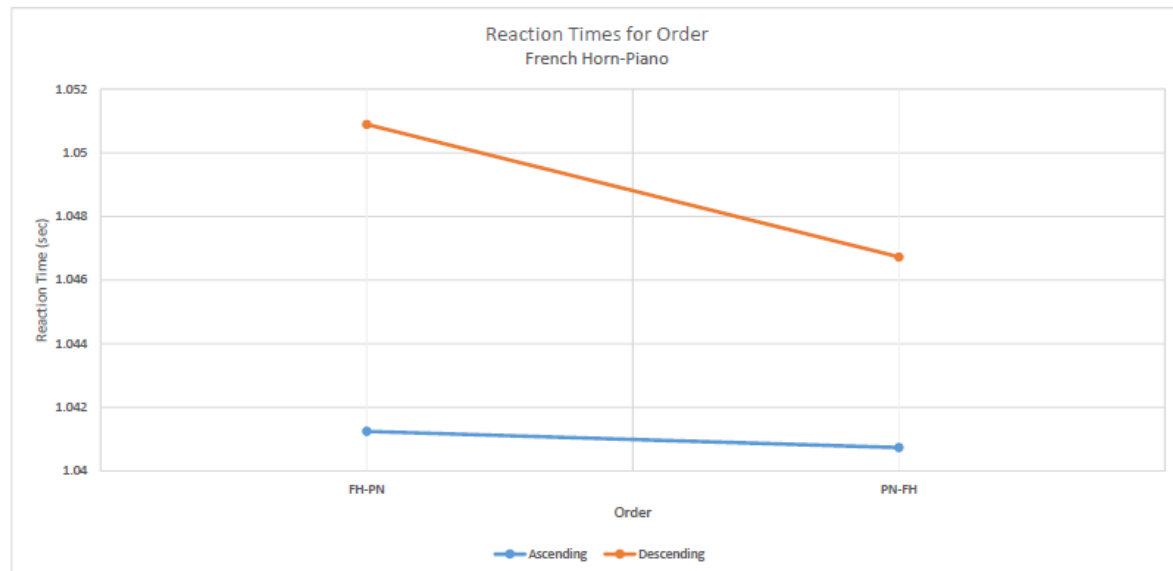
Order	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
MB-VN	.668	.040	.584	.752
VN-MB	.605	.042	.517	.692

Table 2.16: Mean Response Times by Order for BC-MT

Order	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
BC-MT	.631	.042	.543	.718
MT-BC	.645	.042	.558	.731

Table 2.17: Mean Response Times in seconds for all timbre pairs

Order	Mean (sec)
FH-PN	1.819
PN-FH	1.872
MB-VN	1.95
VN-MB	1.83
BC-MT	1.879
MT-BC	1.905

**Fig. 2.8.** Response Times (seconds) for Order, pair FH-PN

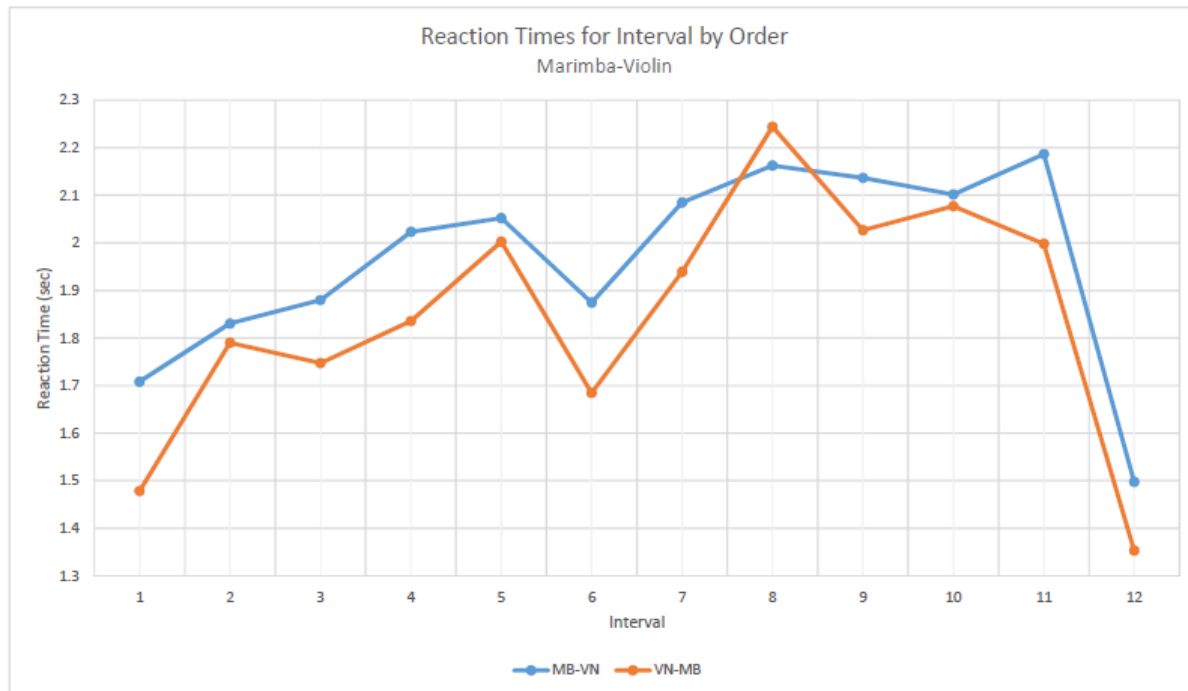


Fig. 2.9. Response Times for Intervals by Order for MB-VN

D. Discussion

The first hypothesis stated that the timbre pair that varied along the dimension of spectral characteristics (MB-VN) should interfere with interval identification more than for other timbre pairs that varied along other dimensions. The results showed that this was in fact not the case as the piano baseline trials did not outperform any other timbre pair in accuracy or response time, and the spectral pair (MB-VN) did not consistently interfere with interval identification more than any other timbre pair. Strangely, the VN-MB and FH-PN pairs (temporal dimension) were found to be the fastest timbre pairs, whereas the MB-VN pair was the slowest. Paired-samples t-tests (including the Bonferroni-Holm correction) were conducted for the differing order between MB-VN and VN-MB for each interval revealed that only the minor second and octave had significantly different mean response times, both of which had some of the highest accuracy and lowest response times overall. While the marimba and violin timbre pair did not interfere with

interval identification (in accuracy or response times) more than any other timbre pair, there were some other curious findings concerning this timbre pair in the response errors. The MB-VN and VN-MB orderings both had the only errors on unisons, most of which were miscategorized as octaves. The spread of miscategorizations also increased for the MB-VN timbre pair overall, and was considerably larger for unison, minor second and minor sixth based on the root-mean-squared errors computed for each timbre pair for each interval (see Table 2.9 and Figure 2.6). The piano baseline in fact demonstrated the smallest spread of errors compared to all the other timbre intervals, with the marimba-violin pair having the greatest spread.

The second hypothesis stated that pitch intervals that were poorly identified in the baseline task would be more prone to interference with timbre. The results from the piano baseline trials were consistent with the previous literature (Killam, *et al.*, 1975; Samplanski 2005). They demonstrated that minor sixths and major sevenths were the most difficult to identify and were also the slowest in response time. There was, however, an interaction between direction and interval not found in these studies, with ascending intervals being more accurate and also faster to identify than descending ones. Paired t-tests revealed (with a Bonferroni-Holm correction) a significant accuracy difference for the major seventh only, and response time differences for the major sixth, perfect fourth, major second, minor third, octave, major third, and minor second. The results indicated that poorly identified intervals were not susceptible to more interaction with timbre. Timbre did not affect accuracy scores on any interval, and although timbre pair did interact with interval in the response time analysis, no systematic effects were seen on poorly identified intervals (minor sixth or seventh). It was suspected that this interaction of timbre pair and interval was in fact a result of the MB-VN (spectral) pair, so the repeated measures-ANOVA was completed on Timbre Pair(6)*Direction(2)*Interval, excluding the MB-

VN pair. The interaction between timbre pair and interval was no longer significant, $F(14.0, 293.4) = 1.41$, $\epsilon = .254$, $p = .148$, $\eta_p^2 = .063$. No three-way interactions were found in either accuracy or response time data, indicating no systematic interactions of pitch, timbre and direction.

The final hypothesis involved demonstrating an interval illusion through decreased accuracy scores with an increase in miscategorizations in a specific response category (e.g., an increase in errors for the perfect fifth, with more errors found in the minor sixth category for a congruent interval in the spectral MB-VN pair). Although timbre pair did not affect interval accuracy, as mentioned above, some interesting findings in response errors were found for the marimba-violin pair, which varied in terms of spectral characteristics, particularly for the unison errors found in the MB-VN and VN-MB trials. Robinson (1993) notes that change along spectral centroid, particularly an increase in spectral centroid from the first pitch to the second, can result in an octave error. This may suggest the importance of timbre in the perception of tessitura or register, further demonstrating possible interactions between timbral brightness and pitch height. This phenomena also resembles the interval illusion investigated by Russo & Thompson (2005b). If an illusion was to occur, we would expect to see an increase in unisons being labelled as octaves for the MB-VN pair (increase in spectral centroid), and a decrease in unisons being labelled as octaves in the VN-MB pair (decrease in spectral centroid). The same could be found with octaves, although including the directional component, congruent octaves (descending MB-VN) being more likely labelled as unisons, and incongruent octaves (descending VN-MB) less likely to be labelled as octaves. For the unisons, we do in fact see more octave confusions for the MB-VN unisons, while the VN-MB unisons have one response as an octave and one response as a minor second. The octave errors also change between these two timbre pairs. The congruent

octave (descending MB-VN) contains unisons errors, while the incongruent octave (descending VN-MB) does not contain any unison miscategorizations. The number of miscategorizations of this kind, however, is extremely small (unisons were 97% accurate in both cases), although it is interesting that the only unison errors were made in the marimba-violin timbre pair. No other interval was found to have evidence of interval illusion. In fact, never was a perfect fourth or perfect fifth miscategorized as a tritone. Note that subjective ratings of interval size showed, for example, that an ascending perfect fifth going from a brighter to a duller timbre, could be perceived as smaller than a tritone going from a dull to bright timbre (Russo & Thompson 2005). While the spectral timbre pair did not systematically demonstrate evidence of interval illusion across intervals, the unison-octave confusions found in this timbre condition alone suggest that interval illusion in octave-unison confusion could still be present in highly trained musicians, even though identification of these intervals was among the highest accuracies and fastest response times. This indicates that ease of interval discriminability may not play a large role pitch-timbre interactions.

Another curious finding includes the interactions of French Horn-Piano order and direction in both accuracy and response times, although these effects were only marginally significant. This finding shows that intervals are more accurate when the French Horn is the bottom note, whether in the ascending or descending direction (see Figure 2.4). This finding is replicated in the descending direction only in response times, indicating that response times were faster in the descending direction when the French Horn was on the bottom. The reason for this finding is unclear at this point. One could hypothesize that it could be a result of the difference of attack and decay between the two sounds. Support for this is seen in the overall effect of order on response times, FH-PN being faster than PN-FH. The French Horn has a sloped attack with a

sustained sound, while the piano has a sharp attack and steeper decay. In the FH-PN ordering, the sustain of the French horn leads directly to the sharp attack of the piano, while the PN-FH order has the sharp decay of the piano followed by the sloped attack of the French Horn, possibly contributing to slower response times of interval identification. The interaction of Order*Direction for this timbre pair, however, does not support this finding as increased accuracy and response times are found when French Horn is both first and second (as long as it is on the bottom). The reason for this finding is therefore still unclear.

2.4 Conclusions and Future Directions

Overall, changes in timbre did not affect accuracy scores for highly trained musicians with verified interval identification abilities. Although changes in response times were seen as a result of timbre pair, no systematic interference of timbre pair with interval identification was found. Interesting effects were found however in miscategorizations for the spectral timbre pair, indicating that spectral characteristics might interfere with interval identification more than other timbral dimensions, such as attack time. These miscategorizations occurred with unisons and octaves, intervals that nonetheless had some of the highest accuracy scores and lowest response times. This result indicates that the ability to discriminate and label intervals may not be the determining factor in timbral interference with pitch. To better understand the effects of different timbral dimensions on interval identification, synthesized stimuli could be used in future experiments in order to better control the differing dimensions of the timbres used. For this study, acoustic timbres were used to facilitate applications of the results to music-theoretic discourse.

There are also some key differences between this study and others that have come before that may have contributed to these interesting findings. One primary difference is the type of task used. Previous studies have primarily used *Garner Classification* tasks to investigate perceptual interactions between pitch and timbre (Melara & Marks, 1990a, b, c; Krumhansl & Iverson, 1992; Pitt, 1994, Allen & Oxenham, 2014), as well as direct comparisons and subjective ratings (Beal, 1985; Russo & Thompson, 2005b). These tasks do not explicitly access musicians' categorical knowledge of interval labels, which could possibly have led to an increase in the salience of timbre. This experiment required participants to access their categorical knowledge of intervals, thus encouraging them to hear through timbral differences in order to identify the intervals, leading to a possible lessening of the effect of timbre. The task used here may have also forced participants to rely more heavily on the temporal (periodicity) cues of pitch chroma (important for pitch-interval perception) rather than tonotopic spectral cues (or pitch height), limiting the possible interactions between pitch and timbre. The subjective ratings used in Russo & Thompson (2005b), for example, could have been based more on spectral cues (thus related more to pitch height), causing participant responses to be more susceptible to interference with timbre.

Another key difference is the population of musicians used. No previous studies tested the interval identification abilities of their musicians, and simply assumed that those with musical training possessed the ability to consistently identify the intervals being used in the experiment. The screening procedure, however, demonstrated that many musicians have difficulty in interval identification, as ten of out 31 participants scored below 18/26 on the screening test. There was also a great variety of performance within the participants that passed the screening procedure, with accuracy scores ranging between 60% and 97% for the full

experiment. This demonstrates that interval discrimination varied quite a bit within trained musicians, and that we should therefore be extremely careful in selecting participants when investigating phenomena that involve interval discrimination. Because the participant population used in this study was screened for pitch-interval accuracy, and overall had high accuracy scores and fast response times for interval identification, the effect of timbre could have been lessened due to their high degree of skill. More interference effects could be seen with a population less adept at interval identification. Future investigations could involve several distinct musician populations with varying degrees of interval discrimination.

3 Theoretical Investigation

3.1 Introduction

The previous chapters have demonstrated some interesting interactions that can occur between pitch and timbre. The interactions shown, however, have occurred in laboratory situations which are highly controlled, making it very difficult to speculate about how these interactions might occur in more complicated, real musical situations. For music theorists, how these phenomena transpire in music is a primary concern. Does timbre interact with pitch in such a way as to change our perception of pitch structure on a larger scale? Do timbral modifications affect any other musical phenomenon, such as formal boundaries or perception of motivic content?

Unfortunately, the current experimental work cannot answer these questions directly due to its limited scope, but we can speculate on these matters. Fortunately, there is other psychological work that has dealt with the function of pitch and timbre in more complicated scenarios, particularly the work of Albert Bregman and others on auditory scene analysis. Bregman's work primarily investigates the factors that contribute to the perception of auditory streams, of which pitch and timbre both play, arguably, equally important roles.

Klangfarbenmelodien is therefore a perfect real-world musical phenomena to investigate the role of pitch and timbre as it features a relatively similar importance of both. This section will therefore focus on the aspects of pitch and particularly timbre which contribute to hearings of unbroken linear (or sequential) Weberian style *Klangfarbenmmelodien* in the chamber music of Carter and Webern, using the principles of auditory scene analysis. The effect of timbre on formal and sectional boundaries will also be discussed. Re-orchestration will be used as a tool in

order to better investigate the effect that timbre has on the creation of a single perceptual line, as well as creation of formal boundaries. Before the analysis can begin, I will outline some of the primary principles of auditory stream segregation as discussed by Albert Bregman, and will identify which aspects are important for the perception of *Klangfarbenmelodie*.

3.2 Auditory Scene Analysis

Auditory scene analysis is a theory that focuses primarily on how the auditory system determines whether a sequence of acoustic events results from either one or multiple sources (McAdams & Bregman, 1979). If one source is perceived, then a single integrated line is heard (integration), whereas if multiple sources are perceived, then multiple segregated lines are heard (segregation). These “auditory streams” are mental representations formed from the physical acoustic sequences, which we will see are often perceptually flexible and can be heard as either integrated or segregated under various conditions. Bregman outlines two types of stream segregation, one which he calls *primitive* segregation (based on evolution and biology), and *schema-based* segregation (based on learned patterns, these are susceptible to effects of attention) (Bregman, 1990, chapter 1). Because *Klangfarbenmelodie* typically features atonal musical pitch material, learned schemata (such as tonal syntax) do not typically apply (Bregman, 1990, chapter 5). For the purposes of this project, I will focus exclusively on the role of *primitive* segregation. There are two types of primitive segregation discussed by Bregman: *sequential integration* (how events are mentally combined into one sequential line), and *vertical integration* (how simultaneous events are mentally fused into a single entity) (Bregman, 1990, chapters 3 & 4 respectively). Because the analytical concern here is that of horizontal-style Weberian *Klangfarbenmelodie*, I will only be discussing *sequential integration*. There are several factors that affect segregation

and integration (to be discussed), all of which are very much contextually dependent, and often depend on the listeners' attention. For the purpose of this chapter, I will discuss each factor separately, even though they often compete with one another in more complicated acoustic scenarios (McAdams & Bregman, 1979).

Sequential integration (how events are combined into one sequential line) is governed by two primary features: (1) frequency separation (or distance of pitch between the two tones) and (2) the speed of the sequence. Pitches that are more similar (i.e., close together, around five semitones or less) are more likely to integrate into a single line, while pitches that are far apart are more likely to segregate into multiple auditory streams (Bregman, 1990, chapter 3). The boundaries between integrated and segregated lines caused by pitch and temporal distances of the tones are called the 1) *temporal coherence boundary* (where pitch differences are too large to hear as one coherent line) and 2) the *fission boundary* (where two interleaved sequences of events are too close to hear as separate streams). Generally, the faster the sequence, and the wider the pitch range, the more likely the sequences are to segregate into separate streams (McAdams & Bregman, 1979). These boundaries are often not clear-cut, however, as both frequency separation and tempo contribute to stream segregation. Therefore, the wider the pitch distances are, the more likely segregation will be perceived *even* at slower tempos that typically promote integration under conditions with less frequency separation. Similarly, segregation can also be perceived if the tempo is fast enough, even if frequency separation is quite narrow. Because of this, there is what McAdams and Bregman (1979) refer to as an “ambiguous region” where either integration or segregation can be heard with cognitive effort. Indeed, most music falls in this ambiguous region. As we will see, this region plays an important role in the perceptual analyses of *Klangfarbenmelodie*, particularly because several of the

Klangfarbenmelodien for discussion have wide frequency separations and slow tempi, as well as timbre change, resulting in many musical factors fluctuating simultaneously.

Frequency separation and tempo are not the only factors that contribute, however. The time from the offset of one tone to the onset of the next (termed interstimulus interval) is extremely important for sequential streaming of two alternating tones (Bregman, Ahad, Crum, & O'Reilly, 2000). The shorter this duration, the more likely two tones (in the same frequency range, as this depends on pitch as well) are to form separate streams, even if the duration between the two tones is negative because the tones overlap temporally (Bregman *et al.*, 2000).

Another feature related to offset-to-onset duration is tone regularity (e.g., ABABA versus AAABBAABA). Bregman briefly discusses regularity of tone sequences as a phenomena that should greatly affect streaming, but cites experimental research that has shown that streams formed by a primitive process are not affected by the predictability of regular sequences (Bregman, 1990, chapter 8). A few recent studies have examined the effect of tone regularity on streaming, often leading to null or contradicting results (Handel, Weaver, & Lawson, 1983; Rogers & Bregman, 1993; Bendixen, Denham, Gyimesi, & Winkler, 2010; Andreou, Kashino, & Chait, 2011; Snyder & Weintraub, 2011). Generally, tone regularity is thought to be a secondary parameter that has a stabilizing effect in stream segregation (Bendixen *et al.*, 2010), which is highly dependent on context (Handel *et al.*, 1983). I argue that tone regularity is extremely important in the perception of the *Klangfarbenmelodie* to be discussed below, as the regularity of alternation helps to provide some perceptual stability in sequences in which the time between pitch onsets varies irregularly. This predictability leads to greater sequential integration, as tone irregularity (when combined with pitch and temporal irregularity) causes the streams to gain perceptual independence, causing greater segregation. Pitch materials, tempo, temporal

construction, and regular tone alternations therefore play an extremely important role in sequential integration of *Klangfarbenmelodie*. If the frequency and onset conditions are constructed by the composer in such a way to induce segregation rather than integration, a coherent *Klangfarbenmelodie* will not be perceived. One will be more likely to hear two or more instruments playing in counterpoint rather than being integrated into one perceptual line.

Two other very important factors that influence streaming are timbre and loudness. Loudness has been shown to be subordinate to frequency separation in its effect on streaming, whereas timbre is more complicated (Bregman, 1990, chapter 3). As the current study investigates, there are common acoustic features of pitch and timbre (such as spectral components), which create issues in discussing factors that cause auditory stream segregation. McAdams & Bregman (1979) cite timbral brightness (which has been shown to overlap with the perception of pitch height) as a key dimension of timbre that can affect stream segregation. Generally, similar timbres will promote sequential integration, while dissimilar timbres are more likely to promote sequential segregation. Timbre change has in fact been shown to impact streaming as much as moderate changes in frequency separation (Moore & Gockel, 2002), indicating that both pitch and timbre play an important role in stream segregation and integration. The use of timbre and loudness in *Klangfarbenmelodie* is therefore extremely important. A composer must orchestrate carefully to ensure sequential integration is possible between timbres that are not too dissimilar. Dynamics must also be accounted for. As we will see in the subsequent analyses, timbral similarity and carefully chosen dynamic contrasts are very important aspects of what allows for the perception of sequential *Klangfarbenmelodien*.

The following analyses of *Klangfarbenmelodien* will therefore focus on the features of timbre and pitch structure that promote the hearing of a single melodic line, as well as the

promotion of certain sectional and formal boundaries. These analyses will be based on the principles of auditory stream segregation discussed above, emphasizing the roles of frequency separation (pitch and register), tempo, onsets and tone alternations, timbral similarity and dynamics. To further investigate the effect of timbre on streaming, the effect of re-orchestration will be explored for each excerpt. I will begin with examples of simple two-voice textures, and then progress to more complex, multi-voice compositions, ending with a discussion of timbre's effect on sectional and phrase boundaries in a complex, nine-voice texture.

3.3 Carter Duets and *Klangfarbenmelodie*

Carter is not typically a composer cited for his use of *Klangfarbenmelodie*, although he started making use of the technique in *Eight Etudes and a Fantasy* in 1950 (Schiff, 1998). Several of his later chamber works feature clear juxtapositions of two instruments, often in a manner which encourages the hearing of a single melodic line that shifts in timbre over time. The works for discussion here are *Esprit Rude/Esprit Doux* (1985) for flute and B-flat clarinet, *Rigmarole* (1996) for bass clarinet and cello, and *Au Quai* (2002) for bassoon and viola. Each of these late works by Carter features a limited pitch vocabulary, and particularly an extensive use of the all-interval tetrachords [0137] and [0146]¹⁵ (Hopkins & Link, 2002; Straus, 2009). These all-interval tetrachords ensure a high level of variety of melodic intervals in these works, creating many unique and novel combinations. These duets, particularly *Esprit Rude/Esprit Doux*, use a technique defined by David Schiff as *stratification*, where the musical texture is divided into separate layers with contrasting harmonies, tone colours, rhythms and expressive characters

¹⁵ This notation refers to the pitch classes (from 0-11) present in "prime form." Prime form is a method for organizing pitch structures in their most simplified format with all pitch classes as tightly packed (or close together) as possible, and transposed (or moved) to always start on pitch class 0 (or C). [0137] is therefore the collection comprised out of a minor second, minor third and perfect fifth (C-C#-E-flat and G).

(Schiff, 1998). While these later duets do feature many passages where the two instruments segregate out from one another, there are also contrasting sections where the two timbres combine sequentially (and even vertically) to form a single melodic (or fused) line, morphing in timbre over time.

Esprit Rude/Esprit Doux

One of the earliest works to feature these types of compositional techniques by Carter is *Esprit Rude/Esprit Doux* (1985). This work, written for flute and B-flat clarinet, was composed to celebrate the sixtieth birthday of Pierre Boulez (Truniger, 1998). The flute and clarinet are often stratified (separated) from one another, which is encouraged by the fact that they are given separate pitch materials (Schiff, 1998). The flute plays the minor third, major third, perfect fourth, tritone, and minor and major sevenths, while the clarinet plays the minor and major seconds, tritone, perfect fifth, and minor and major sixths. The work features a large-scale, slow moving polyrhythm between the flute and clarinet in a 21:25 ratio (Truniger, 1998). As a result, the two instruments rarely have simultaneous attacks, which helps to prevent fusion of the two parts into one, simultaneous (vertical) entity. This is supported by Coulembier's analysis of the rhythmic structure, and his assertion that the use of polyrhythm is primarily to help create distinct identities for the two instruments (Coulembier, 2009). In terms of timbre, these sounds are from the same instrument family, meaning that they are more similar in sound than those from different families. The clarinet (depending on register) is slightly more dark and hollow sounding than the flute, with a slightly more sloped attack compared to the flute's sharper attack onset, while the flute is brighter and slightly noisier. If we look at our 3-dimensional timbre space created from similarity ratings, we can see that they are fairly close together (see Figure 3).

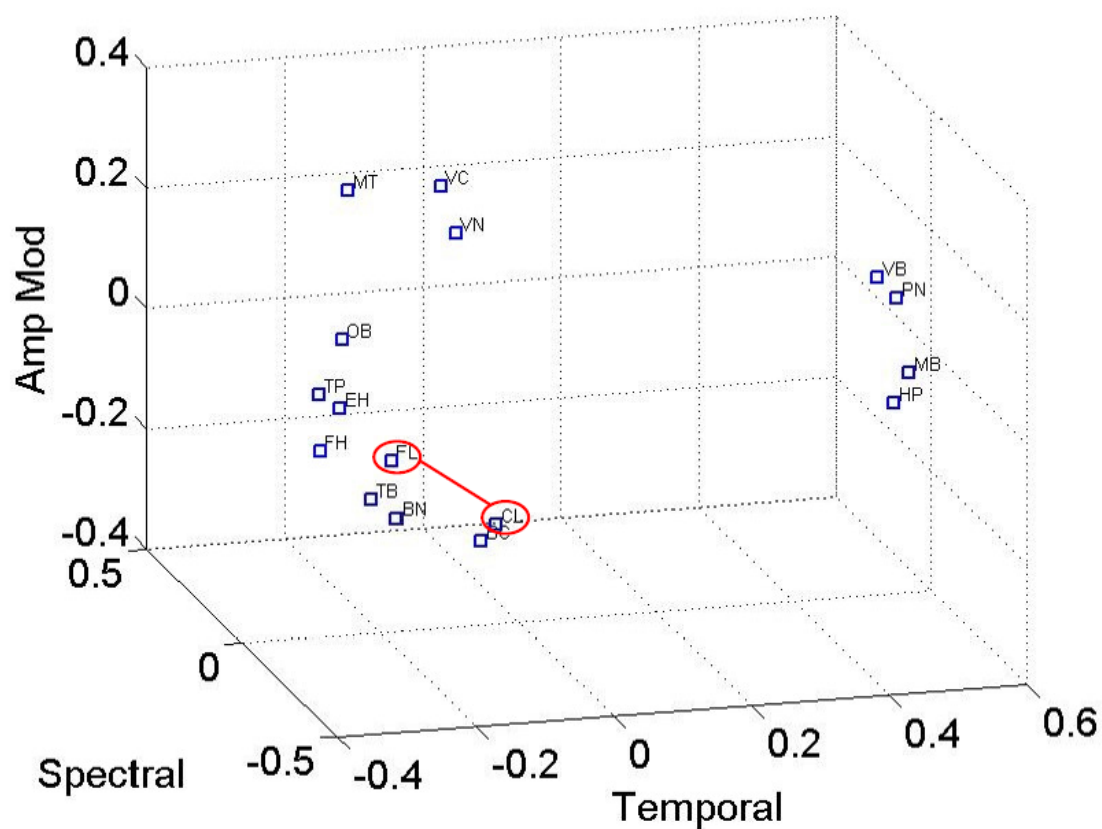


Fig. 3. Timbre Space, similarity of flute and clarinet

The first example we will look at is the introductory gesture in measures 1-4 (see Example 1.1, piano arrangement). This example demonstrates how two instruments similar in timbre, with no temporal overlapping, can combine to form one continuous melodic line. Because the excerpt features minimal temporal overlapping of the instruments, the two instruments combine easily into a single line due to the consecutive attack onsets. The opening gesture is meant to spell out “Boulez,” (B-flat-C-A-E, as “B-ut-la-E”), which is also featured at the end of the piece. The overall pitch organization of this excerpt is typical of Carter during this period¹⁶. Two all-interval [0137] tetrachords are used with exact pitch repetitions in two bar

¹⁶ Many of Carter’s later works feature extensive use of all-interval tetrachords, which has been described by many theorists and Carter himself (see Tuniger, 1998; Schiff, 1998). As a result, set-class analyses for the Carter examples cited here focus specifically on locating all-interval tetrachords where possible.

groupings. As a result, the excerpt is quite symmetrical: the beginning and ending occur on B-flat 5, with the exact pitch repetitions of C4 in the flute, A3 in the clarinet, and E5 in both the flute and clarinet. The excerpt also closely follows the principles outlined in auditory scene analysis that promote sequential integration. The offset-to-onset times between tones are short, if not immediate, and the tempo is slow enough to promote sequential integration between the instruments. The registers are much wider in spots than is specified by Bregman (five semitones), but this is compensated for by the slow tempo (McAdams & Bregman, 1979), timbral similarity (many of the compound leaps are within the same instrument) and also by some dynamics and performance factors (to be discussed). These help to integrate the line sequentially in spots where one might expect large pitch leaps to cause segregation (see Example 1.2, combined version of original flute and clarinet parts).

Example 1.1. *Esprit Rude/Esprit Doux*, mm 1-4

The musical score for Example 1.1 shows measures 1-4 of *Esprit Rude/Esprit Doux*. The score is for Piano, Flute (Fl.), and Clarinet (Cl.). The time signature is 3/4. The Flute part is in the right hand, and the Clarinet part is in the left hand. The Piano part is in the left hand. The Flute part starts with a forte (f) dynamic and a marcato marking. The Clarinet part starts with a piano (p) dynamic and an espress. marking. Both parts feature triplet markings (3) and a [0137] bracket. The Piano part ends with a marcato marking and a dynamic change from f to ff.

Esprit Rude/Esprit Doux by Elliott Carter
 © Copyright 1985 by Hendon Music, Inc.
 Reprinted by permission.

Example 1.2. *Esprit Rude/Esprit Doux*, mm 1-4, combined version

The musical score for Example 1.2, *Esprit Rude/Esprit Doux*, mm 1-4, combined version, is presented in a single staff. The key signature is one flat (B-flat), and the time signature is 3/4. The score is written for Flute (Fl.) and Clarinet (Cl.). The Flute part begins with a forte (*f*) dynamic, followed by a piano (*p*) dynamic, and ends with a fortissimo (*ff*) dynamic. The Clarinet part starts with a piano (*p*) dynamic, followed by a fortissimo (*ff*) dynamic. The score includes various musical notations such as triplets, slurs, and dynamic markings (*f*, *ff*, *p*, *poco*). The Flute part has a triplet of eighth notes in measure 1, a slur over measures 2 and 3, and a triplet of eighth notes in measure 4. The Clarinet part has a slur over measures 2 and 3, and a triplet of eighth notes in measure 4. The score is marked with *poco* in measures 2 and 3, indicating a slight change in tempo or dynamics.

Esprit Rude/Esprit Doux by Elliott Carter
© Copyright 1985 by Hendon Music, Inc.
Reprinted by permission.

Along with timbral similarity and lack of temporal overlap, another feature which contributes to the hearing of one continuous line is the consistent, two-by-two timbral alternation of timbres (i.e., the flute has two notes, followed by two clarinet notes, etc.). This helps to provide predictability to the change of timbre, which provides stability across the wide pitch fluctuations and changes in tone duration. The regular alternation also prevents each instrument statement from being heard as a separate independent event, unconnected from the previous statement. This type of independence will be discussed shortly. Registral and dynamic markings also aid in the hearing of one perceptual line, such as the low, darker flute sound on the C4 in measure 2 leads more easily into the very dark low clarinet A3 in the same measure. The crescendo in the clarinet line makes the transition from the low flute C4 much cleaner, and the performer in the recording treats this entry with great care, coming in dynamically quite under the flute in order to grow out of the entrance and become more audible (Carter, *Nouvel Ensemble Moderne*, 2000). Despite the repetitions of intervals and pitch classes throughout the excerpt, the timbral alternations do not always align with these repetitions. This is particularly true in the middle of the excerpt where there are several interval repetitions (at the exact same pitch levels). For example, the ascending compound perfect fifth in measure 2 is within the clarinet, whereas the same compound perfect fifth in descending form is between the flute and

clarinet in measure 4. These timbre alternations overlaid on the repetitive pitch structure help to create variation within tightly organized pitch constructions. They also affect the sectional and phrase structure of the excerpt (see Example 1.3).

Example 1.3. *Esprit Rude/Esprit Doux*, mm 1-4, phrase divisions shown by the dotted line

The musical score for measures 1-4 of *Esprit Rude/Esprit Doux* is shown. The score is in 2/4 time and features a flute (Fl.) and a clarinet (Cl.). The flute part starts with a forte (f) dynamic and a triplet of eighth notes. The clarinet part enters in measure 2 with a piano (p) dynamic and a triplet of eighth notes. The flute part has a triplet of eighth notes in measure 3. The clarinet part has a triplet of eighth notes in measure 4. The score is divided into two sections by a dotted line. The first section is labeled [0137] and the second section is labeled [0137] and 'Echo'. The dynamics are f, p, f, and ff. The tempo is marked 'poco'.

Esprit Rude/Esprit Doux by Elliott Carter
 © Copyright 1985 by Hendon Music, Inc.
 Reprinted by permission.

Example 1.3 shows the phrase construction of the excerpt, which is divided into two sections, shown by the dotted line. The pitch structure, organized into two [0137] all-interval tetrachords, do not align with this boundary, even though the second flute gesture in measure 3 has a unique character. This second flute gesture functions like an echo figure of the clarinet in measure 2, even though the pitch and interval repetitions are not identical. The hearing of the echo figure is in fact aided by the timbre change from clarinet to flute, and is supported by the change in dynamic from *forte* to *piano*. After this, the only part of the line which sections off from the main line is the clarinet entry in measure 4, within the descending compound perfect fifth from the flute E5 to clarinet A3. The register shift is very large, and the timbral shift from bright flute (flutter tongued) to low, dark, heavily articulated and *forte* clarinet A3 helps to separate them, causing a clear break, whereas the previous instrument alternations combine sequentially quite well.

In order to better understand how much of these properties of the excerpt are due to timbral changes, I created a timbre neutral version (piano arrangement) of measures 1-4 (see Example 1.1). These, and all subsequent arrangements were created using Finale 2007. The piano version of this excerpt has many of the same qualities of the original, although some key aspects of the segmentation and phrase structure are quite different. Firstly, there are no issues in hearing a continuous, sequentially connected line. The use of timbres from the same instrument across the parts greatly reduces any effect that large register leaps had on any potential segregation heard in the original, indicating the importance that timbre indeed does play in stream segregation. The phrase separation observed in the original version with the clarinet gesture in measure 4 can still be heard with some effort in the piano arrangement, but is much less distinct. The register leap and dynamic shift still allow for a phrase separation in measures 2 and 4, but this separation is less perceptible. This results in another noticeable change, which is the lack of call-response effect, making the echo figure in measure 3-4 very difficult to hear. One could imagine that this echo could be improved with the use of exaggerated dynamic shift, but is inherently less echoic in nature without the timbre shift present.

The opening of *Esprit Rude/Espirit Doux* provides a clear example of how two separate instrumental lines can combine into one, single *Klangfarbenmelodie*. Timbral similarity, consistent timbral alternation, relative registral proximity and lack of temporal overlap all aid the sequential integration of these two instruments in the first example. Many *Klangfarbenmelodien* in the music of Carter are not so straightforward however. As with many musical textures, overlapping of instrumental lines inevitably occurs, causing more difficulty in hearing a single monophonic *Klangfarbenmelodie*. There is therefore a higher likelihood of hearing what David Schiff refers to as *stratification* (Schiff, 1998). The following excerpt will demonstrate that

hearing a single line in these instances is still possible, although the boundary between sequential integration and segregation is rather fuzzy. Similarly to the opening excerpt, sequential integration can be heard due to timbral similarity, registral proximity, limited temporal overlapping, and consistent alternation of timbres. In striving to hear a single *Klangfarbenmelodie* between the flute and clarinet, the complex relationship between the interleaving parts becomes more apparent, allowing for an elevated dialogue to be perceived between the two lines not demonstrated in the upcoming timbral-neutral version of the same excerpt.

The final 6 measures of *Esprit Rude/Espirit Doux* demonstrate the type of complex relationship possible in single line *Klangfarbenmelodien* that feature moments of slight temporal overlap between parts. The pitch materials are similar to the opening, with the “Boulez” motive present in measures 87-88, and many of the gestures, as well as pitch and interval content, are alike, but are re-contextualized (see Example 2.1). All-interval tetrachords [0137] are used in measures 83 and 84, with a [0134] in measure 85, followed by two more all-interval tetrachords in measures 86-88, [0146] and [0137]. Unlike the first excerpt, however, more extensive temporal overlapping is observed between the two lines, creating the possibility to perceive a contrapuntal, stratified texture, rather than one continuous line.

Example 2.1. Carter *Esprit Rude/Esprit Doux*, mm 83-88

The musical score for Example 2.1, Carter *Esprit Rude/Esprit Doux*, measures 83-88, is presented in two systems. The first system (measures 83-85) features a Flute (Fl.) and Clarinet (Cl.) part. The Flute part starts with a *pp* dynamic and includes a triplet of eighth notes. The Clarinet part starts with a *f* dynamic and includes a triplet of eighth notes. The second system (measures 86-88) features a Piano part. The Piano part starts with a *pp* dynamic and includes a triplet of eighth notes. The score includes dynamic markings (*pp*, *f*, *ff*) and articulation (accents, slurs). Interval brackets are labeled with numbers in brackets: [0137], [0137], [0134], [0146], and [0137].

Esprit Rude/Esprit Doux by Elliott Carter
 © Copyright 1985 by Hendon Music, Inc.
 Reprinted by permission.

Clear temporal overlapping is seen in Example 2.1. The registral spacing between the two parts shifts between small and large distances throughout the excerpt, creating areas where segregation is more likely (such as the compound minor third in measures 85 and 87). These are typically preceded and/or succeeded by smaller intervals within the octave, however, limiting the segregating effect. Although the parts overlap temporally, providing an increased possibility of segregation, the onsets and timbral alternations are evenly distributed, just as in the previous excerpt. The instruments alternate in a one-by-one fashion rather than in twos, and there is only one instance of two pitches with the same timbre in a row, clarinet E4 to clarinet Eb4 (measures 84-85). From measure 83 until this point, the alternation had been FL-CL. This alternation, however, switches to CL-FL through measures 85-88 after the clarinet plays briefly disrupts the

consistent alternation in measures 84-85. The proximity of the overlapping pitches also helps to create one perceptual line, in that they are close enough rhythmically to be sequentially integrated, but do not overlap long enough to promote segregation. This results in the combined melody seen in Example 2.2. Also, adjacent notes in the same timbre are often separated by rests, making a within-instrument connection more difficult to hear. There are a few spots where the same instrument slurs into its next pitch however, such as the flute at the opening of the excerpt (A5 to E4). Because of the proximity of the pitches within the same instrument, and the direct connection without break or change of articulation, it could be possible to hear those two pitches (flute A5-E4) as part of the same line with a clarinet B-flat 5 in counterpoint against it (see measure 83, Example 2.1). These instances of direct connection within the same instrument provide some issues when attempting to hear one continuous line (see measures 84 CL, 85 FL, 86 CL, and 87 both FL and CL). The overall timbral similarity and consistent timbral alternations combat within-timbre connections, and help to promote between-instrument sequential integration. In terms of phrase structure, the excerpt seems to be divided into three main sections: measures 83-84 (framed by a restatement of the descending compound perfect fourth), 85-86, and 87-88 (see Example 2.2). These segments mostly line up with the boundaries of tetrachords used in the excerpt (see Example 2.1).

Example 2.2. *Esprit Rude/Esprit Doux*, mm 83-88, combined version with phrase boundaries

Esprit Rude/Esprit Doux by Elliott Carter
 © Copyright 1985 by Hendon Music, Inc.
 Reprinted by permission.

In the piano arrangement of measures 83-88, these phrase boundaries are not nearly as apparent. The only audible boundary in the piano arrangement is at the end of measure 84 (see Example 2.1), which seems to occur due to the large rest before the following entries in measure 85. The lack of timbral diversity also greatly reduces tendencies for segregation, even where there is large frequency separation, such as in measures 86 and 87. Hearing segregated parts in these areas is still possible, but requires some cognitive effort in the timbral-neutral version. Similarly to the previous excerpt, the lack of timbral variety also dissipates the “call-response” dialogue between the parts heard in the original.

To demonstrate that timbral similarity cannot always effect the formation of a single line between two instruments, I will briefly examine the following excerpt from measures 32-35. The original and re-orchestrated in timbre-neutral versions do not vary significantly. The large registral separation plays an important role in segregating the two lines, as well as the alternation of the onsets and timbres between the lines, which unlike many of the previous examples, is not even, alternating FL-CL-CL-FL-CL-FL-CL-CL-CL-FL-CL-CL. This unpredictable alternation of

timbres helps to separate out the lines, increasing the independence of each instrument. This is also aided by the fact that the within-timbre onsets are closer than the between-timbre onsets, thus promoting stream segregation as opposed to sequential integration of a single merged line.

Example 3. *Esprit Rude/Esprit Doux*, mm 32-35

The musical score for Example 3, *Esprit Rude/Esprit Doux*, measures 32-35, is presented for Piano, Flute (Fl.), and Clarinet (Cl.). The Piano part begins with a forte (*ff*) dynamic and features a series of chords and single notes. The Flute part has a melodic line with various articulations, including accents and slurs. The Clarinet part also has a melodic line with various articulations, including accents and slurs. The score is in 4/4 time and includes a key signature of one flat.

Esprit Rude/Esprit Doux by Elliott Carter
 © Copyright 1985 by Hendon Music, Inc.
 Reprinted by permission.

Carter and *Klangfarbenmelodie*: Examples with Timbral Dissimilarity

The following excerpts feature timbral alternations that are between instruments of different families, winds and strings. The wind instruments featured are bassoon and bass clarinet and the strings are cello and viola (which is not present in our 3D space, but which would probably be between VN and VC). These sounds generally vary most along the dimension of attack (strings bowed), as well as in the spectral dimension (see 3D timbre space, Figure 4). The excerpts discussed will only be those in which little to no temporal overlap occurs in order to focus more closely on the features of timbre that allow for sequential integration to take place.

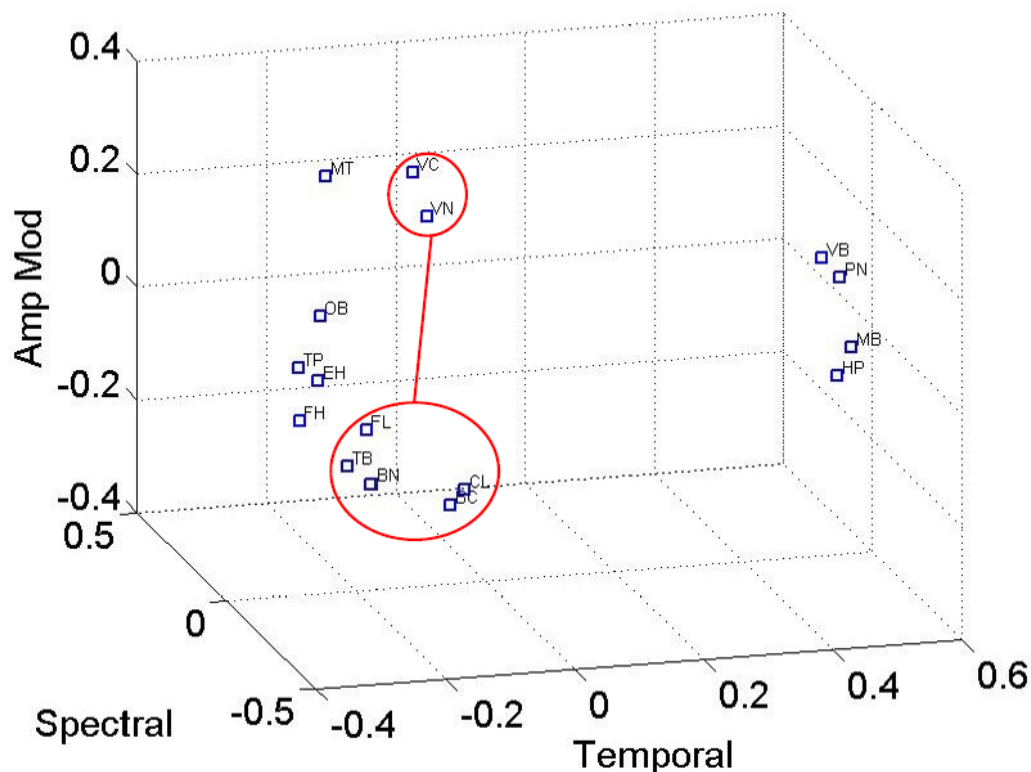


Fig. 4. 3D timbre space, winds vs. strings

The first example with dissimilar timbres is from Carter's *Rigmarole* (1996), written for cello and bass clarinet (see Example 4.1). This excerpt features little to no temporal overlapping of the parts, and like Examples 1 and 2 from *Esprit Rude/Esprit Doux*, contains even alternations of onsets and timbres, easily permitting a sequential hearing of the combined melodic line. The pitch construction is similar to Carter's other works of this period: a predominant use of all-interval tetrachords. All-interval tetrachords are used at the beginning and end, with other tetrachords and trichords in the middle. The excerpt goes from [0146], [0137] to [0358], [0247], then trichords [025], [027], [048], back to the tetrachords [0148] and [0137]. Interestingly, the middle section that diverges from using all-interval tetrachords uses sets that progressively decline in variety of interval content. The first tetrachord that is not an all-interval tetrachord is

[0358] in measure 19 with the interval vector $\langle 012120^{17} \rangle$, followed by [0247] which has the interval vector $\langle 021120 \rangle$. The trichords that follow slowly decrease in variety of interval content: [025] with $\langle 011010 \rangle$, [027] with $\langle 010020 \rangle$ and finally [048] with $\langle 000300 \rangle$. This is followed by the [0148] tetrachord with the interval vector $\langle 101310 \rangle$, and the final all-interval tetrachord [0137]. This seems to have a streamlining effect in that the interval content becomes more restricted, before it goes back to being unrestricted with the all-interval tetrachords.

Example 4.1. *Rigmarole*, mm 15-25

The musical score for 'Rigmarole' by Elliott Carter, measures 15-25, is presented in three systems. The first system (measures 15-19) shows a piano part with complex rhythmic patterns, including triplets. Interval vectors [0146], [0136], and [0358] are labeled above the first three measures. The second system (measures 20-22) continues the piano part, with interval vectors [0247], [025], [027], and [048] labeled above the measures. The third system (measures 23-25) shows the piano part concluding with interval vectors [0148] and [0137] labeled above the measures. Performance instructions include 'Bc.', 'Vc. (pizz)', '(Segregated)', and '(arco)'.

Rigmarole by Elliott Carter
 © Copyright 1996 by Hendon Music, Inc.
 Reprinted by permission.

¹⁷ An interval vector is a representation of the interval content present in a given group of pitches, going from one to six semitones from left to right. $\langle 123456 \rangle$ Each number in the vector represents the quantity of the interval type. Intervals larger than a tritone, e.g., a perfect fifth, are inverted to be represented as smaller than a tritone (e.g., perfect fifth would be inverted to a perfect fourth). For example, the vector $\langle 012120 \rangle$ shows that there are no semitones or major sevenths, one major second or minor seventh, two minor thirds or major sixths, one major third or minor sixth, two perfect fourths or perfect fifths, and no tritones.

Unlike the excerpts from *Esprit Rude/Esprit Doux*, Carter uses two instruments that are from different families, and are further away in timbre space (see Figure 4). In our timbre space, bowed cello and bass clarinet are not terribly far apart along the spectral dimension, nor the temporal dimension, but do differ in amplitude modulation. Because most of the cello notes in this example are played pizzicato, we could speculate that it would be closer in the timbre space to the percussive sounds (marimba, harp), making the bass clarinet and cello in *Rigmarole* in this section more distant along the temporal, or attack time dimension. However, as the bass clarinet is also staccato in *Rigmarole*, its attack is more like the cello, reducing this dissimilarity effect. Despite the increased timbral dissimilarity, hearing one conceptual line here is rather simple, with the exception of one spot which seems to segregate into two lines in measure 22 (see Example 4.1). Even though the lines do not overlap in time (as the cello notes are pizzicato), each part streams separately in measure 22 because within-instrument connections seem to be easier to hear than across- or between-instrument connections. This is likely due to the increase in tempo (more tones/second), as well as the register shift, which allows for segregation on the basis of timbre. The cello plays two pitches in a row before the bass clarinet enters, possibly encouraging the segregation of the two lines by encouraging within-timbre connection instead of across-timbre connections. We will return to examine this section in the timbre-neutral version to see if orchestration is the primary reason for the segregation here.

What seems to be more evident in the examples with dissimilar timbres compared to the *Esprit* excerpts is an increased sense of dialogue between the two instruments. This sense of dialogue is rather interesting as it provides distinction between the two parts even while they both combine to form a *Klangfarbenmelodie*. In *Rigmarole*, this dialogue or “call-response” feature is enhanced by the grouping alternations of the timbres and pitch materials and their

systematic change over the course of the excerpt. For example, at the beginning of the excerpt in measure 15, the groupings change from two notes in each instrument to three and then to one (back and forth). Just like the echo figure in Example 1.1 (measures 1-4 of *Esprit Rude/Esprit Doux*), the timbral alternations help to provide a higher level dialogue across pitch materials that are not identical or necessarily related. The pitch patterns in this excerpt (Example 4.1) are not equal (the same interval distance), nor are they even going in the same directions. The timbre alterations help to emphasize the grouping structure and patterning and help to relate materials on the surface that would not necessarily be related had orchestrational changes not been present.

If we compare these to the timbre-neutral piano version, some of these very important attributes that defined the original excerpt are lost. While it is far easier to integrate the lines into a single line, the dialogue or “call-response” feature almost completely vanishes. While the dynamic changes help to address this issue, and further performance factors could help this (such as exaggerating the dynamic contrasts, time delays, etc.), the missing timbral difference between the instruments, most notably along the attack dimension reduce the call-response effect drastically. The grouping structure at the beginning of the excerpt is also completely gone in the piano version (two-three-one). Interestingly, the more contrapuntal section in measure 22 that was clearly separated in the original remains separated here, suggesting that the temporal, registral, and rhythmic features of this bar are more responsible for the separation of the two lines rather than timbral differences.

Example 4.2. *Rigmarole*, mm 15-25, piano (timbre neutral)

15

Piano

[0146] [0136] [0358]

20

[0247] [025] [027] [048]

(Segregated)

23

[0148] [0137]

sf *f* *f* *f*

mf *mf* *f* *mf* *f*

p *mf* *f* *p* *f*

mf *mf* *f* *f* *f*

p *f* *f* *p* *f*

Rigmarole by Elliott Carter
© Copyright 1996 by Hendon Music, Inc.
Reprinted by permission.

While the timbral neutral version eliminates many of the features of the original, what about a version with two different instruments which are closer in timbre space? To examine what would happen with more timbral similarity, I created a version that includes the original bass clarinet with a substituted bassoon (as featured in *Au Quai*, to be discussed). In our timbre space, these two instruments are the closest instruments, while remaining distinct from one another (clarinet was closer to bass clarinet, but their timbres are nearly identical). In this version, we gain back the call-response feature lost in the piano version, as well as the grouping

structure at the opening. The tendency toward segregation of the two parts is also less evident than the original, and is more similar to the piano in that the parts blend more easily into one conceptual line.

Another example of *Klangfarbenmelodie* which contains two instruments from different families is *Au Quai*; the bassoon and viola (see Example 5.1). These two instruments, similar to bass clarinet and cello from *Rigmarole*, differ along amplitude modulation more than spectral centroid or temporal factors. Similarly in this example, the viola plays pizzicato the majority of the time, although interestingly alternates more often between arco and pizz than the cello in *Rigmarole*, which was primarily pizzicato. The bassoon and strings in our timbre space, again, do not differ greatly along spectral or temporal dimensions, although vary greatly the most along the dimension of amplitude modulation. The bassoon often compensates for the dissimilarity of attack by playing staccato, as was seen in *Rigmarole* by the bass clarinet.

Example 5.1. *Au Quai*, mm 12-17, perceived monophonic line

Au Quai by Elliott Carter
 © Copyright 2002 by Hendon Music, Inc.
 Reprinted by permission.

We would imagine that the pizzicato viola (similar to the cello), would vary more greatly on the temporal dimension compared to bassoon. Similarly to the bass clarinet in *Rigmarole*, the staccato bassoon limits the dissimilarity to the pizzicato viola. Carter plays with this dynamic in the excerpt to aid in the integration of one line, and also enhance the call response feature present. Like the other excerpts, mostly all-interval tetrachords are used alternating with other tetrachords: [0146] in measure 12, [0248] in measure 13, [0146] in measure 14, [0126] in measures 15-16, and [0167] in measure 17 (see Example 5.1). One interesting timbral effect on pitch in this excerpt is the major second between the viola G4 and bassoon A4, which sounds much larger than it is (especially compared to the timbre-neutral version). This might have to do with the playing effort involved in the bassoon compared to the viola. Like the excerpt from *Rigmarole*, the call-response feature here is aided by the consistent alternations in timbre (generally one-to-one), and is further emphasized by articulation matching between the bassoon and viola, which is particularly evident in measures 16-17. The short-long bassoon gesture is imitated by the viola in measure 17 using a pizz-to-arco playing technique. These timbral alternations once again help to relate pitch content that would not necessarily be related. The interval sizes in the groupings are not identical, and interval directionality is not always the same across groups. Overall, it is possible to hear one *Klangfarbenmelodie* that also contains some internal independence of the parts through a call-response dialogue provided by timbral difference.

In the piano, or timbre-neutral, version, the grouping structures and call-response features are lost (see Example 5.2). The connection between the G4 and A4 in measures 14-15 seems much more fluid, and has no “stretching effect.” The phrase groupings also change as a result, as timbre is no longer a factor in grouping. The segments are at the end of measure 12 and in the

middle of measure 16, the first of which cuts through a previously grouped line within the bassoon.

Example 5.2. *Au Quai* mm 12-17, piano (timbre neutral) phrase boundaries

Au Quai by Elliott Carter
© Copyright 2002 by Hendon Music, Inc.
Reprinted by permission.

Conclusions about Carter

After having examined the two-voice *Klangfarbenmelodie* from several of Carter's duets, we can come to several conclusions about what timbre can add to two-voice textures, as well as what issues we should be aware of moving on to pieces with more lines and more timbres. From the above analyses, confirmed by Albert Bregman's research in auditory stream segregation, we can see that timbral similarity leads to greater ease in hearing sequential integration, while timbral dissimilarity promotes segregation into two lines in counterpoint. This is evidenced by the increase in ease in hearing one conceptual line in the timbre-neutral versions rather than the original, multi-timbre versions.

Sequential integration is very much dependent on musical factors (tempo, rhythm, register), as can be seen by several sections in the excerpts from *Esprit Rude/Esprit Doux* and *Rigmarole* which differed little between the original and timbre-neutral versions. Another factor that affects the integration of one line is the regularity of timbral alternation. Even or consistent alternation seems to aid the integration (one-to-one, two-to-two), whereas unpredictable alternations (seen in *Esprit*, Example 3) lead to greater independence of the parts, and thus more segregation. Lastly, timbre also affects surface grouping and phrase separations. This is most clearly seen in *Rigmarole* and *Au Quai* with the clear call-response feature leading to pitch-interval groupings. In the timbre-neutral versions, these qualities were almost completely lost. In the more timbrally similar version of *Rigmarole* with bassoon and bass clarinet, the call-response feature is kept, although is slightly less apparent than in the original.

3.4 Webern and *Klangfarbenmelodie* in Quartet, op. 22 and Concerto, op. 24

Not all music, or all *Klangfarbenmelodien* for that matter, occurs in just two lines. Webern's *Klangfarbenmelodie*, as discussed earlier, is far more complicated, featuring complex alternations of instruments, often in a pointillistic manner. Webern's chamber music often uses traditional musical forms, such as canon, sonata, variation, and rondo (Whittwall, 2008). Both of the works to be discussed are in traditional forms: movement II of the Quartet, op. 22 for B-flat clarinet, violin, tenor sax and piano is in rondo form, while movement II of the Concerto for Nine Instruments, op. 24, for flute, oboe, B-flat clarinet, trumpet, French horn, tenor trombone, violin, viola and piano, is in ternary form. The purpose will be to extend the discussion begun in the Carter examples to works that contain more instruments and more complex textures. These excerpts contrast with one another in texture and type of *Klangfarbenmelodie*. The excerpt from

Quartet, op. 22 features a complex layering in which timbre helps to create different types of musical textures (such as melody and accompaniment, monophony, counterpoint, etc.) that are not present when timbral differences are removed. In Webern's Concerto, op. 24, the effect of timbre change on phrase and formal divisions will be discussed.

Quartet, op. 22, movement II

Similarly to the some of the Carter examples, Webern's Quartet, op. 22, movement II includes temporal overlapping lines and instrumentation. With the added complexity of multiple timbres, is it still possible to hear a single line in this multi-instrument example? What would the conceptual line look like for this excerpt? I argue that it is indeed possible to perceive and follow a single line in this example, and that this is often made possible through timbre. The movement as a whole is composed in rondo form, and moves away from Webern's typical horizontal symmetries and tightly organized patterns featured in works before this one (Bailey, 1991). The form, ABACADA is slightly unclear according to Bailey, supposedly roughly based on the scherzo from Beethoven's piano sonata, op. 14 (Bailey, 1991). The excerpt for discussion, measures 69-88 (see Example 6.1), is within section C of the rondo form, and is constructed entirely out of untransposed rows (Bailey, 1991). Within measures 69-88, the role of the instrumentation often changes, causing the instrumental lines to switch between varying texture types including traditional monophonic *Klangfarbenmelodie*, and multiple hierarchical layers forming melody and accompaniment. Given the instrumentation, one might expect that the tenor saxophone and B-flat clarinet would be more likely to integrate into a single line, while the violin and piano would likely segregate out from the texture as they are more dissimilar. This is not always the case, although we will see that the violin (when bowed) often functions in

counterpoint or as an echo, because it is more easily segregated out. I shall begin by showing the sections of the excerpt which are primarily monophonic, followed by those that are homophonic in texture.

Example 6.1. Quartet, op. 22, mvt II, mm 68-88, original

The musical score is divided into three systems. The first system (measures 68-75) shows the Violin (Vn.) and Clarinet (Cl.) playing a melodic line, with the Piano (Pn.) providing harmonic support. The second system (measures 76-84) continues the melodic development, with the Piano part becoming more active. The third system (measures 85-88) shows the Violin and Clarinet playing a melodic line, with the Piano part providing harmonic support. The score includes various dynamics (ff, p, sfz, pp) and articulations (pizz, arco).

Anton Webern „Quartett für Geige, Klarinette, Tenorsaxophon und Klavier [op. 22“
© Copyright 1932 by Universal Edition A.G., Wien/UE 10050

There are two sections within this excerpt where a continuous monophonic *Klangfarbenmelodie* can be heard: measures 75-78 and 81-84, where measures 79-80 functions as a type of piano interjection not connected to the material before or after. Measures 75-78

feature little to no temporal overlap of gestures, making the perception of one continuous line more straightforward than measures 80-83 (see Example 6.2).

Example 6.2. *Quartet* op. 22, mm 75-78, monophonic line



Anton Webern „Quartet|für Geige, Klarinette, Tenorsaxophon und Klavier|op. 22“
© Copyright 1932 by Universal Edition A.G., Wien/UE 10050

Example 6.2 shows the perceived monophonic line of the excerpt, while the original version can be viewed in Example 6.1. While some of the tones in measures 75-78 overlap temporally, the staccato and tenuto staccatos help to shorten the notes and aid their connection, as does the slow tempo. The consistent alternation of timbres between the tenor saxophone and clarinet also aids in the creation of a monophonic line, and helps to attenuate the segregating effect of the large register leaps, particularly in measures 75-76. The violin enters in measure 77 with a variation of the clarinet gesture. Because of its timbral dissimilarity to the clarinet, and also because it breaks the consistent alternation of timbres up to this point, this gesture does not function as a continuation of the line, but rather as an echo figure. The clarinet C5 in measure 77 instead connects to the clarinet G4 in measure 78. The clarinetist in the Ensemble Intercontemporain recording helps this between-clarinet connection (Webern, Ensemble Intercontemporain, 2000) by exaggerating the length of the pitch in measure 77 to better connect to the clarinet note in measure 78, thus enabling the hearing of one continuous line. This line is completed by the final pizzicato violin D4, which is well connected to the clarinet that preceded it. It is interesting that this connection between the clarinet and violin in measure 78 is possible, while the bowed violin gesture in measure 77 does not connect well with the line preceding it.

The violin pizzicato in measure 78 is far duller in brightness here, and it is performed slightly lower in volume than the preceding clarinet note, allowing it to more easily match the duller clarinet and tenor saxophone sounds that precede it. The pizzicato articulation helps the string instrument sound more like the staccato wind instruments, and less like a stringed instrument. The arrangement that I created for piano combines easily into one monophonic line, and the echo figure played originally by the violin acts much more as a continuation of the line than as an echo of the previous statement (see Example 6.1, mm 75-78). These two versions of measures 75-78 demonstrate the importance of timbral similarity for the creation of a horizontal *Klangfarbenmelodie*. None of the perceptual issues present in the original are present in the piano-only version, although the sense of dialogue and expressive variation created by timbre change is completely lost.

The second section which clearly features a single monophonic *Klangfarbenmelodie* line is slightly more complicated, as part of the line breaks temporarily into two lines in melody and accompaniment and features some vertical fusion. Example 6.3 (measures 81-84), shows the perceived monophonic line which is heard, where Example 6.1 shows that many of these measures contain other musical material which separates out into melody and accompaniment (mm 81) or fuses into blend (mm 84).

Example 6.3. *Quartet* op. 22, measures 81-84, monophonic line

Anton Webern „Quartet|für Geige, Klarinette, Tenorsaxophon und Klavier|op. 22“
© Copyright 1932 by Universal Edition A.G., Wien/UE 10050

The line begins on the tenor saxophone G#4 and continues down to the A3 in the saxophone. Meanwhile, the piano E4 and violin pizzicato B-flat form a (brief) supportive accompaniment for the saxophone melody in measures 81-82 (see Example 6.1). While the musical figures in the tenor, piano and violin do not differ in rhythm, it is clear that the tenor sax is in the foreground here, creating two hierarchical layers. This tenor saxophone A3 in measure 82 is accompanied by another violin gesture, similar to the echo in measure 77, which is not integrated with the tenor, but is on a similar higher layer of segregation (melody) over the piano (accompaniment, see Example 6.1). The tenor A3 then moves to the piano note B-flat 3 in measure 83, which enters at roughly the same volume, and in a register where the piano timbre connects well to the tenor. This line continues from here as a single monophonic line that does not segregate into multiple parts. The piano B-flat 3 is followed by the clarinet E4, which moves downward to the piano B4, fusing vertically with the clarinet F4. The clarinet performer makes quite a decrescendo here, allowing for the vertical blend to occur, which causes the clarinet note to integrate sequentially with the closer pitch in the piano (B4) as opposed to the clarinet F4 in measure 84 (Webern, *Ensemble Intercontemporain*, 2000). This piano line finishes the short section by moving to the C#5 grace note, with the G4 following it (quickly) nearly being fused vertically with the C#5. This monophonic line is also aided by similarities in articulation, as each instrument that participates in the single monophonic line is sustained and contains no staccatos or accents.

It is clear that the original version proves some difficulty for the formation of a sequentially integrated line. Timbral and registral similarity help to integrate the melodic line, while timbral and registral dissimilarity help to create segregated hierarchical layers of (brief) melody and accompaniment. This is evidenced by the separation of the violin and piano from the

tenor sax in measures 81-82. The inherent dissimilarity of instruments can, however, be compensated for with register and dynamics, as was shown by the connection of the tenor saxophone and piano in measures 82-83. The use of articulation also aids the sequential integration of instruments, as all of the instruments in the perceived line are sustained without staccato in measures 80-84. The piano arrangement of this section fares much differently (see Example 6.1, mm 81-84). Most of the excerpt has much less definition into distinct lines, with the beginning sounding homophonic, while measures 82-84 mostly fuse together into a single line. I am not able to follow the initial G#4 down to the A3, only the highest pitches stand out as a “melody” (G#4 across measures 81-82, E5 and B4 in measures 83-84). The dialogue feature is also lost between the parts.

Now that we have looked at the sections with traditional monophonic *Klangfarbenmelodie*, we can discuss the sections which segregate into two hierarchical layers, forming homophonic textures. These sections are from measures 70-72 and 84-87, with measures 73-74 forming another piano gesture, disconnected from measure 72. The first section at measure 70 begins with the accompaniment line, in the piano (F4-B3, see Example 6.4). The violin enters at B-flat 4 with a pizzicato articulation, which connects to the tenor saxophone C#5. This is followed by the same “echo gesture” that has been previously seen in the other examples in the violin. This time, the saxophone timbre connects easily to the clarinet timbre to form a single line. The clarinet line continues, while the saxophone drops down two octaves to become part of the accompaniment line. We do not follow the saxophone line from the C#5 to D2 because the register leap is too large, and the higher and brighter clarinet line stays in the foreground. The tenor instead follows the piano B3 from measure 71. This section does not continue onward to measure 73 because much like in the monophonic section from measures 75-78, the bowed

violin gesture separates from the clarinet gesture, leading into a piano figure (as seen in measures 79-80). The piano arrangement of this section, similarly to the other sections, lacks definition of the different parts, although a homophonic texture can still be heard.

Example 6.4. *Quartet* op. 22, mm 70-72

Anton Webern „Quartet[t]für Geige, Klarinette, Tenorsaxophon und Klavier[op. 22“
© Copyright 1932 by Universal Edition A.G., Wien/UE 10050

The final section which segregates into two lines in melody and accompaniment is from measures 85-88 (see Example 6.5). Here the timbral material does little to differentiate the lines, the long half-note values easily segregate out from the combined grace note gestures played by the clarinet and violin, which function as a kind of melody. The final bar fuses vertically into one entity. The piano version does not differ greatly from the original: a clear two-part texture is still present, although of course the specific instrument blends are not.

Example 6.5. *Quartet* op. 22, measures 85-88

Anton Webern „Quartet[t]für Geige, Klarinette, Tenorsaxophon und Klavier[op. 22“
© Copyright 1932 by Universal Edition A.G., Wien/UE 10050

***Quartet op. 22* Conclusions**

Webern's op. 22 demonstrates how more than two instruments can combine to form a single monophonic *Klangfarbenmelodie*. This work also shows how timbre can contribute to the creation of more complex textures, such as melody and accompaniment layers both containing timbral shifts across time. Timbral similarity helps to sequentially integrate monophonic and melodic lines together, while timbral dissimilarity seems to be responsible for the segregation of homophonic layers, particularly in measures 70-72. Performance factors, such as modified pitch length, dynamics, and modified articulation, help to connect the *Klangfarbenmelodie* in the original, multi-instrument version, while the timbre-neutral piano version makes some of the melodic lines nearly impossible to follow (see measures 81-84). Overall, the texture changes create a type of four-part structure where the excerpt begins homophonically in measures 70-72, moves through a piano gesture to monophony in measures 75-78, moves through another piano gesture, to another monophonic section (with some melody and accompaniment) in measures 81-84, and finally to a completely homophonic section in measures 85-88 (see Example 6.6, summary). Not only does this excerpt feature alternating orchestration creating interesting *Klangfarbenmelodien*, these *Klangfarbenmelodien* also shift through several complex texture types. These texture shifts are often only possible to hear because of timbre change and timbral contrasts, as was shown through the lack of definition in the timbre-neutral versions (piano-reduction).

Example 6.6. *Quartet* op. 22, mm 68-88, summary

The musical score is divided into three systems, each illustrating different textures and dynamics:

- System 1 (Measures 68-75):**
 - Measure 68:** Piano (Pn.) plays *ff*. Violin (Vn.) and Clarinet (Cl.) play *p*. Violoncello (Vcl.) is marked (pizz).
 - Measures 69-71:** Labeled **Homophony**. All instruments play in unison.
 - Measures 72-73:** Labeled **(piano gesture)**. Violoncello (Vcl.) plays (arco) and (echo). Violin (Vn.) plays (pizz) and (fusion). Piano (Pn.) plays.
 - Measure 74:** Labeled **(piano gesture)**. Violoncello (Vcl.) plays *sfz*.
- System 2 (Measures 76-84):**
 - Measures 76-78:** Labeled **Monophony**. Violoncello (Vcl.) plays *p*. Violin (Vn.) plays (arco). Clarinet (Cl.) plays (pizz).
 - Measures 79-81:** Labeled **(piano gesture)**. Violoncello (Vcl.) plays *sfz*. Violin (Vn.) plays *p*. Piano (Pn.) plays.
 - Measures 82-84:** Labeled **Monophony (with some homophony)**. Violoncello (Vcl.) plays *sfz*. Violin (Vn.) plays (arco). Clarinet (Cl.) plays *p*. Piano (Pn.) plays.
- System 3 (Measures 85-88):**
 - Measures 85-88:** Labeled **Homophony**. Violoncello (Vcl.) plays *pp*. Violin (Vn.) plays (arco). Clarinet (Cl.) plays. Piano (Pn.) plays.

Anton Webern „Quartet|für Geige, Klarinette, Tenorsaxophon und Klavier|op. 22“
© Copyright 1932 by Universal Edition A.G., Wien/UE 10050

***Klangfarbenmelodie* in Webern's Concerto for Nine Instruments, op. 24, mvt**

II: Timbre's effect on Motivic and Formal Boundaries

We have seen how Carter and Webern have created *Klangfarbenmelodien* that can be heard as one continuous monophonic line, as well as *Klangfarbenmelodien* that switch between monophony and melody and accompaniment. The final piece for discussion will be the second movement of Webern's Concerto, op. 24 for flute, oboe, B-flat clarinet, trumpet, French horn, trombone, violin, viola, and piano. Unlike many of the previous examples, the Concerto features an unambiguous melody (played by the wind, string, and brass instruments), which is accompanied by the piano. Analyses completed by Kathryn Bailey, Christopher Wintle and Leopold Spinner reveal that there are many disagreements regarding the form and structure of the work. My analysis will not focus on when hearing one line is possible (as a melody & accompaniment texture is always present), but will investigate how timbre (and more specifically timbral similarity), either aids or disrupts linear melodic connections, resulting in different motivic and phrase segmentations. Ultimately, the original version of the concerto promotes more surface-level segmentation, which allows for the hearing of a type of sentence form (complete with fragmentation) (Caplin, 1998). In timbre-neutral versions, these segmentations do not occur, and therefore sentence form cannot be perceived. This longer-range type of hearing promoted by the timbre-neutral version reflects why certain analytical decisions might have been made by Spinner, Wintle and Bailey, as timbre and orchestration were not always considered by these theorists.

The three analyses by Bailey, Wintle and Spinner focus heavily on pitch materials, row forms and set classes. Both Spinner and Bailey rarely mention timbre and orchestration, and instead focus nearly exclusively on pitch class, row forms, and motivic analysis for their

discussion of formal construction in the second movement. Spinner uses pitch class repetitions, row forms, and motives (A, B, and C) to delineate the formal boundaries present in the period at the beginning of the second movement (mm 1-28, Spinner, 1955). He refers to the “principal part” as being played by the wind and string instruments, while the accompaniment is played by the piano (Spinner, 1955, pp. 47). After this brief mention of timbre, he simply refers to the melody as the “principal part,” making no mention of the continuous fluctuation of instrumentation that is responsible for the formation of the three motives, which are building blocks of the phrase divisions present in the period. Bailey is equally reluctant to discuss timbre and orchestration as important factors in formal construction. She briefly mentions that the final A section of the movement contains all nine instruments, where the initial A section does not, providing greater timbral variety (Bailey, 1991, pp. 250). Following this, her discussion of form in the initial A section is reduced to a description of row forms present, with nearly no mention of the role of orchestration in the creation of formal divisions.

Wintle does slightly more to mention the importance of instrumentation. In an introductory section, he describes the performance and conducting of the work, emphasizing the contradicting role of the piano reduction. Here he highlights how well the second movement of the concerto fits “under the hands”, and subsequently notes that piano reductions were an important part of Webern’s process, and that he often coached conductors from the piano (Wintle, 1982, pp. 77-78). Wintle then quotes the conductor Hermann Scherchen on the potential limitations of keyboard training alone, which generally causes musicians to decompose melodic relationships into small parts (Wintle, 1982, pp. 78). Wintle recognizes this issue in the second movement of the concerto, stating that “compositional tension in the work resides in the opposition between linear arches...and the fragmentation of the instrumentation...” (Wintle,

1982, pp. 81). To combat this, Wintle again cites Scherchen saying that performers must mentally sing the complete line of which they are a part (Wintle, 1983, pp. 81). This recognition of the issues present in the performance of Webern's *Klangfarbenmelodie* brings much to the analysis of the second movement of the Concerto. Unfortunately the function of timbre and orchestration is not an important aspect of the subsequent analysis. Formal divisions are cited as arising primarily from tempo and dynamic shifts, and the focus of the analysis following the formal discussion is to demonstrate how these divisions are reinforced on a small and large scale by structured pitch materials (Wintle, 1983, pp. 86). As a result, neither Wintle, Bailey nor Spinner fully discuss the importance of timbre and orchestration in the formal construction of the movement. Their analytical disagreements (to be discussed below) often stem from their prioritization of pitch structures over orchestration. As I will demonstrate, orchestration helps to clarify formal divisions that are inherently more difficult to perceive in timbre-neutral versions (particularly in piano reductions).

Formally, each of the three theorists divide the overall form of the work into three large sections in an ABA form. Most of their disagreements of the form are centered on the first A section, which all three authors describe as a large-scale period form (Spinner, 1955; Wintle, 1982; Bailey, 1991). Each of the authors divide the first 28 measures into three sections: a period with both an antecedent and consequent (see Example 7.1), followed by a third section functioning as either a prolongation or a transition. Both Wintle and Spinner place the beginning of their consequent phrase in measure 10 where the pitches from the opening of the piece are repeated (G-E-flat), and the tempo returns to normal after the *calando* marking (Spinner, 1955; Wintle, 1982). Bailey notes, however, that this area can be interpreted either as a new beginning of the consequent phrase or as cadential motion closing the antecedent phrase (Bailey, 1991).

Bailey opts for the second cadential reading, because the underlying row structure supports a new beginning in measure 13 rather than in the second half of measure 10. The theorists also disagree over the exact placement of the end of the consequent. Bailey places the end of the consequent after the final trombone gesture in measure 24 (mirroring her placement at the end of the antecedent phrase in measure 13), with the following measures 25-28 functioning as a transition to the B section in measure 29 (Bailey, 1991). Spinner however places his end of the consequent in measure 22 before the entrance of the trumpet at the *tempo* marking, with measures 23-28 functioning as a prolongation of the consequent (Spinner, 1955). Wintle places the end of his consequent in between Bailey and Spinner in measure 23, with the trumpet entrance marking the beginning of the prolongation (Spinner, 1955; Wintle, 1982).

My own formal analysis of the A section of opus 24, movement two, contains similar large-scale boundaries, although the exact nature of my motivic and smaller-scale formal divisions are rather different. I argue that the original version promotes more segregation of motives and small-scale formal boundaries, whereas timbre-neutral versions promote longer range hearings of the work, which more closely resemble the analyses cited above. I argue that the original version helps to promote a hearing of the form as a compound period constructed out of two Caplinian sentence forms (Caplin, 1998), which is made possible through timbre change. This timbre change determines which musical materials are connected together (forming basic and contrasting ideas through timbral similarity) and which are segregated (forming fragmentation through timbral dissimilarity). Other factors such as frequency separation, distance between onsets, and dynamics also play a role. The importance of timbre for hearing sentential form is demonstrated by the timbre-neutral versions as they promote the hearing of a large-scale period without the internal divisions necessary to hear sentential form.

Example 7.1. Concerto op. 24, mvt II, mm 1-28, analysis summary of Bailey, Wintle and Spinner

L.S: Antecedent
C.W: Antecedent
K.B: Antecedent

Piano

Consequent
Consequent

calando ----- tempo

10

pp p mp p mp

Prolongation

calando tempo Extension Transition calando

19

p mp p f f mp p

28

pp

Anton Webern „Konzert für 9 Instrumente/op. 24“
© Copyright 1948 by Universal Edition A.G., Wien/UE 34118

I will begin by discussing the completely timbre-neutral version scored for piano reduction (see Example 7.2). My own analysis of the timbre-neutral version is very similar to Wintle's analysis. The antecedent phrase spans measures 1-10, while the consequent phrase ends in measure 22 with an extension phrase in measures 23-28. Wintle describes the interior divisions of the antecedent and consequent as being represented by the repetitions of dynamics, that generally move from quiet to louder and back to quiet (such as the *pp* to *mp* back to *pp* in measures 1-5, Wintle, 1982). These dynamic changes do represent a systematic manner of dividing phrase structure, although I argue they are very difficult to hear, particularly in a timbre-neutral piano reduction version. In my own analysis, I have very few internal phrase divisions as the surface segmentation of melody is extremely difficult to hear out when both the melody and accompaniment lines are played by the same instrument. It is often the case that the piano accompaniment (from the original) crosses register with the melodic line, such as in measures 11-12, making the melodic line extremely difficult to hear, even with the dynamic alterations highlighted by Wintle. As a result, I propose that the internal phrase divisions are linked more to the occurrence of rests than melodic groupings, resulting in the groupings seen in Example 7.2 (marked under brackets). In instances where the melodic line is easily segregated from the accompaniment due to register, such as in measures 6-7, rests in the melodic line provide grouping structure. These internal phrase divisions, as we will see, are much longer than those in the original, as well those in a version arranged for solo B-flat clarinet and piano (to be discussed). The piano reduction promotes much longer hearings of phrase segmentations, and therefore finding internal divisions based on pitch or motivic grouping is rather difficult. This helps to explain some of the differences between the analyses of Spinner, Wintle and Bailey, as each theorist's segmentations are based on different factors. Spinner bases his divisions on

motivic repetitions and tempo markings, whereas Wintle makes use of the semi-repetitive dynamic markings. Bailey on the other hand disregards the surface of the music and segments based on row structure, which in the piano reduction, seems as valid as any other type of segmentation, likely because hearing surface groupings are extremely difficult in a timbre-neutral setting.

Example 7.2. Concerto op. 24, mvt II, mm 1-28, piano reduction

Antecedent

Piano

Consequent

calando ----- tempo

Extension

calando tempo calando

28

pp

Anton Webern „Konzert für 9 Instrumente op. 24“
© Copyright 1948 by Universal Edition A.G., Wien/UE 34118

I created another type of timbre-neutral version where the *Klangarbenmelodie* of the original is played by a single voice (clarinet) with piano accompaniment, in order to better segregate the melody and accompaniment from one another (see Examples 9.3-9.5). In this version, melodic grouping structures can be more easily heard, allowing surface repetitions and motivic groupings to be perceived. Because the melodic line is still relatively timbre-neutral, there are fewer internal divisions than can be observed in the original, although a clearer melodic grouping structure is perceived than in the piano reduction simply because the melody can be separated out from the accompaniment. Compared to the piano reduction, in the antecedent phrase of this version the groupings are no longer simply based on the rests and piano groupings, so the first and third groupings are extended (see Example 7.3). The consequent phrase is now a three-part structure instead of the two-part structure based on the rest in measure 18. This consequent phrase also contains an echo figure in the piano which was not heard in the piano reduction (see Example 7.4). As in many of the Carter examples, changes in timbre aid a type of “call-response” dialogue between lines. This feature can be seen here with the echo figure in the extension in measures 21-22 (see Example 7.5). This echo helps to provide the cadential closure needed for the antecedent and consequent phrases, which reinforces the larger formal boundaries, making them much easier to hear in this version compared to the piano reduction. Overall, the melodic grouping structure of this version begins to show the structural building blocks that are required in order to hear sentential form in the original, although it still lacks the appropriate segmentation and fragmentation to be perceived as a sentence.

Example 7.3. Concerto op. 24, mvt II, mm 1-11, Antecedent phrase in Clarinet version

Period: Antecedent

Clarinet in B \flat

Piano

Ant.

B \flat Cl.

Pno.

calando

Echo

Anton Webern „Konzert für 9 Instrumente op. 24“
 © Copyright 1948 by Universal Edition A.G., Wien/UE 34118

Example 7.4. Concerto op. 24, mvt II, mm 11-22, Consequent phrase in Clarinet version

Period: Consequent

B♭ Cl.

Pno.

Cons.

Anton Webern „Konzert für 9 Instrumente op. 24”
© Copyright 1948 by Universal Edition A.G., Wien/UE 34118

Example 7.5. Concerto op. 24, mvt II, mm 23-28, Extension in Clarinet version

B♭ Cl.

Pno.

Echo

calando

Anton Webern „Konzert für 9 Instrumente op. 24”
© Copyright 1948 by Universal Edition A.G., Wien/UE 34118

The original version of the work promotes much more internal segregation due to timbre change within the melody. This allows for the hearing of sentential form within the framework of the compound period observed by Spinner, Wintle and Bailey. I will begin by showing how the addition of timbral information encourages hearing the antecedent phrase as constructed out of a sentence (Example 7.6). Generally, the instrumentation changes each measure, with some extending over two measures. The grouping does not simply align with each change of instrument, however: some timbral changes encourage connections to be made across instrument change, while others promote separation, causing the groupings to change, and not solely to rely on rests and pitch groupings. The basic idea in measures 1-2 is played by the muted trumpet followed by the viola, while the second in measures 4-5 is played by the violin and clarinet. The second basic idea is rhythmically reversed, with the two quarters preceding rather than following the half note. Although the instruments come from different families, the use of quiet dynamics, mutes, and decrescendos helps to connect these basic ideas into single units. Webern in fact aids the connection of the violin (muted) to clarinet in measures 4-5 through the use of decrescendo. In the previous timbre-neutral versions, measures 4-6 were connected together, but here they are played on clarinet followed by flute. These instruments are more similar than clarinet and violin, which connect together in the previous two measures. There is, however, no connection in measures 5-6 because the dark clarinet E-flat at pianissimo does not connect well to the high, bright flute D6 at the considerably louder dynamic of mezzo piano.

The gestures following the second basic idea are minor sixths played by the flute, oboe and lastly violin. In the previous timbre-neutral versions, measures 7-10 were easily connected together as they contain no rests. These gestures are not connected to each other here, but are separated and function as echoes of the initial flute gesture. This creates a very important aspect

Example 7.6. Concerto op. 24, mvt II, mm 1-10, Antecedent phrase constructed as a Sentence

Sen. _____

9 Instru. *p* Vn. *pp* Pno.

105

Following both Spinner and Wintle, my consequent phrase begins in measure 11 beginning on a brass instrument and on the same pitch classes as the antecedent phrase (G-E-flat, see Example 7.7). Although similar to the antecedent phrase, it contains more variation as well as more segmentation of the melody. The timbral shift from muted trombone (instead of muted trumpet) to (unmuted) viola in measures 11-14 makes hearing a complete unit rather difficult. The connection is more difficult to make between trombone and viola than between trumpet and violin (both muted) due to the separation of register and sharp timbral dissimilarity. As a result, the first basic idea is separated into two smaller gestures played by the trombone and viola, with little connection between the two. This is followed by a basic idea which moves from muted trombone to clarinet, which is an easier connection to make due to the shift in dynamics from piano to pianissimo (similar to the second basic idea of the antecedent phrase, violin to clarinet). These both function as basic ideas, however, due to the recurrence of the trombone. The trombone shifts to another timbre (viola or clarinet), creating two basic ideas formulated specifically out of timbre change. These motives form the presentation, which similarly to the first sentence is followed by continuation composed of fragmentation. The same instruments are used as in the continuation phrase of the antecedent, although out of order. Oboe comes first, followed by flute and then violin. Similarly to the beginning of the continuation between the clarinet and flute in the antecedent phrase, the connection between the clarinet and oboe here is hindered by the shift in dynamic (from pianissimo to mezzo piano), and also a shift in register (up a ninth from E4 to F5). This is once again followed by a piano echo of the violin in measures 21-22 at the same pitch level, but is interrupted by a trombone echo.

This leads to the final phrase which begins in measure 23 at the *tempo* marking with the trumpet at forte dynamic (see Example 7.8). This phrase functions as an extension of the

consequent phrase, ending with a more suitable, uninterrupted cadence (see the echo figure). The final gestures function as an extension of the fragmentation seen in the previous phrase, although one connection is possible between the oboe and violin in measures 25-26, forming a “basic idea”-like unit. These two instruments are rather contrasting, but the decay in dynamics from forte to mezzo piano and the motivic prevalence of the half note to two quarter notes also likely aid this connection. In the timbre-neutral versions, the extension of the consequent formed one large unit, whereas in the original, we are able to hear two separate units which refer to previous motivic materials. This demonstrates the importance of timbre change for the perception of phrase segmentation and motivic connections. The phrase comes to a proper close with the piano echoing the flute A5 to clarinet F5 in measure 28, just as in the antecedent phrase in measure 10.

Example 7.7. Concerto op. 24, mvt II, mm 11-22, Consequent phrase constructed as Sentence

Period: Consequent

Sentence

Presentation

11 *tempo* B.I. (ext.) B.I. Frag.

Instru. Trb. Va. Cl. Ob.

p mp p pp mp

Pno.

p mp p pp mp

Cons.

Sent.

Continuation

18 Frag. Frag. Echo Figure

Instru. Fl. Vn. Trb.

mp p mp

Pno.

p mp p

Anton Webern „Konzert für 9 Instrumente/op. 24”
© Copyright 1948 by Universal Edition A.G., Wien/UE 34118

Example 7.8. Concerto op. 24, mvt II, mm 23-28, Extension

23

Instru.

Tr. *f*

Ob. *f*

Vn. *mp*

Cl. *p*

Pno.

f

mp

p

pp

--- tempo Frag.

B.I.

calando Echo Figure

Echo

Anton Webern „Konzert für 9 Instrumente op. 24“
© Copyright 1948 by Universal Edition A.G., Wien/UE 34118

Conclusions from Webern's Concerto op. 24, mvt II

The second movement of Webern's Concerto has little ambiguity as to which lines are supposed to integrate sequentially. The composition is clearly composed for a *Klangfarbenmelodie* line, accompanied by piano. The use of *Klangfarbenmelodie* encourages more separation of motives, and smaller formal divisions within the large-scale period outlined by previous theorists, allowing the hearing of two sentence forms combining to form a compound period. This is primarily due to the fact that the timbral alternation promotes the hearing of a continuation function (fragmentation). The piano reduction and solo clarinet versions promote much longer phrases (indicated by Spinner, Wintle and Bailey) with fewer internal segmentations. The timbre-neutral versions also promote some formal ambiguities as evidenced by these analysts' differing starting points for the consequent and the extension/transition phrase, which occur because priority is given to row structure, tempo markings, on dynamics, without full

consideration of orchestration. In these instances, the original version provides a much clearer indication through instrumentation of where formal boundaries are located (as evidenced by the sentential structure of the compound period), while the piano reduction contains far more ambiguity because the motivic and row structures alone are not enough to clearly delineate form.

3.5 Closing Remarks

This chapter has investigated the role that timbre plays in the perception of different types of *Klangfarbenmelodie*. Close examinations of the duets of Elliott Carter revealed that two instruments, playing different musical materials, can be sequentially integrated if they are timbrally similar, their alternation is fairly consistent, and they only overlap momentarily in time. Excerpts where timbral alternation is not consistent and overlapping occurs creates a situation where the lines are stratified (or segregated apart) as was seen in measure 32-35 (Example 3) of *Esprit Rude/Esprit Doux*, even when the timbres involved are quite similar. The effects of dissimilar timbres on sequential integration of *Klangfarbenmelodie* was shown through the analysis of excerpts from *Rigmarole* and *Au Quai*, which demonstrated how Carter used articulation, imitation and dynamics to increase the timbral similarity between two instruments of differing instrumental families.

When more than two instruments are involved, the analytical question surrounding *Klangfarbenmelodie* is slightly different, particularly when the instruments overlap in time, as was seen in Webern's Quartet op. 22. In these instances, not only is the question of sequential integration versus segregation important, but we must also ask what type of musical texture is created by orchestration. In opus 22, we observed that the musical texture shifted over time. This depended on the structure of the musical material and what material was segregated or integrated

based on orchestration. The excerpt changed rather rapidly from monophonic *Klangfarbenmelodie* to textures with melody and accompaniment. In the timbre-neutral version for piano alone, the sequential melodies and monophonic lines heard in the original version were nearly impossible to hear through the dense musical texture, indicating that timbre plays an extremely large role in the perception of sequential melodic lines in more complex multi-instrument *Klangfarbenmelodien*.

The second movement of Webern's Concerto for nine instruments, however, demonstrated that even when the melodic connections within *Klangfarbenmelodien* are clearly outlined, timbre and orchestration have a large impact on the formal, phrase and motivic boundaries perceived in the music. The formal ambiguities and disagreements witnessed in the previous analytical work of Leopold Spinner, Christopher Wintle, and Kathryn Bailey occurred because orchestration was not sufficiently considered in analysis; instead row structure and pitch repetitions were used as the determining factors for formal divisions. In the original version for 9 instruments, formal boundaries are clearer, and allow for the hearing of a compound period constructed out of two Caplinian sentence forms, complete with fragmentation and cadences (formed by echoes). The timbre-neutral versions for clarinet and piano and piano alone promote longer-range hearing which prevents internal divisions of motives and phrases. This work clearly demonstrates the importance of timbre and orchestration in the analysis of twentieth century music and *Klangfarbenmelodie*: timbre has been shown to greatly affect not only sequential integration, but vertical fusion, musical texture types, and even the formal and motivic boundaries that can be perceived by the listener.

4 Conclusions and Future Directions

This project has examined pitch-timbre interactions in both an experimental study and in music-theoretical analyses of twentieth century *Klangfarbenmelodien*. In the experimental study investigating the effect of timbre-change interval identification, many of the questions asked were left unanswered. The results showed that poorly identified intervals in the baseline condition were not susceptible to more interference with timbre change, suggesting that the hypotheses regarding musical training,¹⁸ might not be accurate. The task used may have indeed affected the interactions between pitch and timbre, as categorical identification of pitch-distances could have caused the musicians to ignore small changes in perceived interval size possibly induced by timbre. Little headway was also made in determining which dimensions of timbre are responsible for interference with pitch. The results showed that the timbre-neutral condition did not outperform any of the timbre-changing conditions, and none of the timbre-changing conditions proved more or less difficult in interval accuracy or response time. However, some interesting interval miscategorizations revealed that change in spectral centroid could induce more octave and unison errors, supporting the hypothesis that brightness interferes with the perception of pitch height. The lack of overall effect of timbre on pitch-interval categorizations, however, indicates that there was a strong effect of the task on interactions between these two musical parameters. Previous work, using subjective-rating tasks (such as Russo & Thompson, 2005b), may have encouraged participants to focus on an evaluation of pitch height (tonotopic),

¹⁸ The hypothesis in question is that enhanced pitch perception in musicians provides advantages over non-musicians in timbre's effect on pitch. Pitch-intervals with poorer accuracy should be more susceptible to timbral interference as musicians are less precise with those intervals.

whereas the current study asked participants to categorically label pitch intervals, emphasizing pitch chroma (temporal fine structure). The evaluation of pitch height plausibly emphasized in subjective rating tasks could have therefore increased timbre's effect on pitch evaluations, where the emphasis on pitch chroma through interval identification likely decreased these interactions. Overall, the role of musical training, the type of listening/task that can lead to pitch-timbre interactions, and what dimensions of timbre interaction with pitch are still unclear. Future experimental work should continue to isolate these issues as much as possible in order to address the effect of task, musical training, and timbral dimensions separately, so that we might better understand the role they each play in this complex perceptual phenomenon.

Within the music-theoretical investigations, use of Albert Bregman's principles of auditory scene analysis allowed for a detailed musical analysis of the roles both pitch and timbre play in the perception of *Klangfarbenmelodien*. Timbral similarity was a large determining factor in whether or not *sequential integration* was possible in a *Klangfarbenmelodie*. Timbral similarity was not the only factor, as frequency separation, tempo, and inter-onset interval were also key factors. Even when frequency separation and inter-onset interval were not ideal for sequential integration, timbral similarity helped to provide horizontal congruity where one might expect segregation to occur. The Carter examples from *Esprit rude/Esprit doux* showed how two similar timbres could combine to form various *Klangfarbenmelodien*, whereas the examples from *Rigmarole* and *Au Quai* showed how pieces with dissimilar timbres could also be integrated sequentially. Each of these works demonstrated that timbral change could also induce the perception of a "call-response" feature, which was completely lost in timbre-neutral arrangements of the excerpt. These results demonstrate that a crucial aspect of Carter's *Klangfarbenmelodie* was lost when timbral change was taken away.

Within the Webern Quartet op. 22 and Concerto for Nine Instruments op. 24, I showed the effect of timbre change on the formation of various musical textures (melody and accompaniment and monophony), as well on sectional, motivic, and phrase boundaries. In Webern's Concerto, the inclusion of timbre change in the analysis showed that it was possible to hear a whole other layer of formal divisions, notably sentence form. Timbral dissimilarity promoted increased surface separation, which allowed for the perception of fragmentation, whereas timbral similarity allowed for the connection of motives into basic ideas. Each of these excerpts demonstrates the importance of considering timbre and orchestration, particularly in twentieth century analysis. Future investigations should include analysis of more complicated textures (greater than nine instruments if possible), the inclusion of the effects of context on the perception of *Klangfarbenmelodie* (as context has been shown to affect stream segregation, Borchert *et al*, 2011; Moore & Gochert, 2002), and the implications of timbre and blend on vertical, Schoenbergian *Klangbarenakkord* compared to Webernian horizontal *Klangfarbenmelodie* (Iverson, 2009).

As we have seen, the experimental research has still left many questions open, and future research should reduce the complexity of the tasks used in order to address individual questions more accurately and efficiently. The theoretical work, on the other hand, requires expansion in scope and context in order to be applicable to larger bodies of repertoire, with the desired result of eventually building a larger theory regarding timbre's role in twentieth-century music and beyond. Here lies the primary source of conflict between the experimental and theoretical approaches. One needs reduction and simplicity in order to address issues of perception. The other requires expansion in order to build larger theories and be applicable to more musical repertoire. This conflict proves a great difficulty when attempting to form a dialogue between the

two forms of research, making the application of results from one field to the other extremely difficult. Each field addresses questions surrounding the perception of music, and while often about similar phenomena (as was the case in this study), both the nature and the goals of the questions are very different. This provides a real difficulty for those who seek to traverse both fields simultaneously. This however should not prevent attempts at applications of experimental research from being made. A direct link between my experimental results and theoretical work was not entirely possible, primarily because the task of interval identification does not entirely reflect the way in which we actively listen to music. The current project did find a link however in auditory scene analysis for the discussion of the roles of both pitch materials and orchestration on *Klangfarbenmelodien*. While investigating the influence of timbre on pitch-interval identification may indeed be too narrow for direct theoretical applications, how timbre interacts with pitch materials on the small and large scale is an extremely important phenomenon to understand for both underlying music perception and more abstract music analysis. Possible future work in this area that combines the two streams of experimental and music-theoretical research could, for example, continue to investigate the effects of context on pitch-timbre interactions investigated in Krumhansl & Iverson (1992, Experiment 2). The role of atonal and tonal pitch materials on the perception of change in a tone sequence could be further investigated, and this experimental paradigm could easily be expanded to probe the effect of timbre and pitch context on segmentation and auditory streaming, which are already more directly applicable to music-theoretic discourse.

As researchers in music cognition and perception, it is necessary for us to continue to extend research in our own discipline. However, collaboration and dialogue between music psychology and music theory is necessary if advancements are to be made in either field. Music-

theoretical research should strive to utilize psychological findings to ground and strengthen claims, while those in experimental psychology should strive to build on their research in order to make it more widely applicable to real world music perception.

Appendix

Table A.1. Confusion Matrix (all trials) for PN-PN

PNPN	0	1	2	3	4	5	6	7	8	9	10	11	12
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.954	0.034	0.011	0	0	0	0	0	0	0	0	0
2	0	0.041	0.93	0.017	0.012	0	0	0	0	0	0	0	0
3	0	0.015	0.008	0.931	0.046	0	0	0	0	0	0	0	0
4	0	0	0.006	0.04	0.891	0.04	0.006	0.006	0.011	0	0	0	0
5	0	0	0	0.006	0.023	0.901	0	0.064	0.006	0	0	0	0
6	0	0	0	0.012	0	0.017	0.925	0.012	0.029	0	0	0.006	0
7	0	0	0	0	0.005	0.075	0	0.901	0	0	0	0	0.019
8	0	0	0	0.006	0.006	0.025	0.061	0.061	0.718	0.067	0.031	0.018	0.006
9	0	0	0	0.006	0	0.012	0.012	0.006	0.124	0.776	0.041	0.018	0.006
10	0	0	0	0.006	0	0.006	0.043	0.037	0.123	0.043	0.698	0.043	0
11	0	0	0	0	0	0	0.049	0	0.037	0.024	0.116	0.774	0
12	0.011	0	0	0	0	0	0.006	0.023	0.006	0.011	0.006	0.006	0.931

Table A.2. Confusion Matrix by direction (of semitones) for PN-PN

PNPN	0	1	2	3	4	5	6	7	8	9	10	11	12
-12	0.011	0	0	0	0	0	0.011	0.011	0.011	0.011	0.011	0.011	0.92
-11	0	0	0	0	0	0	0.06	0	0.048	0.036	0.133	0.723	0
-10	0	0	0	0.013	0	0.013	0.075	0.025	0.15	0.062	0.6	0.062	0
-9	0	0	0	0.012	0	0	0.024	0	0.167	0.702	0.06	0.036	0
-8	0	0	0	0.012	0.012	0.037	0.085	0.024	0.72	0.061	0.037	0.012	0
-7	0	0	0	0	0	0.028	0	0.934	0	0	0	0	0.038
-6	0	0	0	0.023	0	0.011	0.931	0.011	0.023	0	0	0	0
-5	0	0	0	0.012	0.035	0.835	0	0.106	0.012	0	0	0	0
-4	0	0	0	0.08	0.864	0.034	0	0.011	0.011	0	0	0	0
-3	0	0	0	0.954	0.046	0	0	0	0	0	0	0	0
-2	0	0.058	0.907	0.023	0.012	0	0	0	0	0	0	0	0
-1	0	0.954	0.034	0.011	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.955	0.034	0.011	0	0	0	0	0	0	0	0	0
2	0	0.023	0.953	0.012	0.012	0	0	0	0	0	0	0	0
3	0	0.031	0.015	0.908	0.046	0	0	0	0	0	0	0	0
4	0	0	0.011	0	0.92	0.046	0.011	0	0.011	0	0	0	0
5	0	0	0	0	0.011	0.966	0	0.023	0	0	0	0	0
6	0	0	0	0	0	0.023	0.919	0.012	0.035	0	0	0.012	0
7	0	0	0	0	0.009	0.121	0	0.869	0	0	0	0	0
8	0	0	0	0	0	0.012	0.037	0.099	0.716	0.074	0.025	0.025	0.012
9	0	0	0	0	0	0.023	0	0.012	0.081	0.849	0.023	0	0.012
10	0	0	0	0	0	0	0.012	0.049	0.098	0.024	0.793	0.024	0
11	0	0	0	0	0	0	0.037	0	0.025	0.012	0.099	0.827	0
12	0.011	0	0	0	0	0	0	0.034	0	0.011	0	0	0.943

Table A.3. Confusion Matrix (all trials) for FH-PN

FHPN	0	1	2	3	4	5	6	7	8	9	10	11	12
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.988	0.006	0	0	0	0	0	0	0	0	0.006	0
2	0	0.006	0.959	0.012	0.018	0	0	0.006	0	0	0	0	0
3	0	0	0.008	0.907	0.062	0.008	0.008	0	0.008	0	0	0	0
4	0	0	0	0.018	0.924	0.018	0.006	0.006	0	0.023	0	0.006	0
5	0	0	0	0	0.023	0.878	0	0.081	0	0	0.006	0	0.012
6	0	0	0	0.006	0	0	0.908	0.011	0.017	0.006	0.023	0.029	0
7	0	0	0	0	0.005	0.032	0	0.917	0.014	0.005	0	0	0.028
8	0	0	0	0.012	0.006	0.006	0.042	0.018	0.786	0.065	0.065	0	0
9	0	0	0	0.006	0.006	0	0.012	0.018	0.089	0.798	0.036	0.036	0
10	0	0	0.006	0.006	0	0.006	0.031	0	0.123	0.031	0.716	0.074	0.006
11	0	0	0	0.006	0	0	0.095	0.012	0.024	0.012	0.107	0.738	0.006
12	0.011	0	0	0	0	0	0	0.029	0	0.006	0	0.011	0.943

Table A.4. Confusion Matrix by direction (of semitones) for FH-PN

FHPN	0	1	2	3	4	5	6	7	8	9	10	11	12
-12	0.011	0	0	0	0	0	0	0.045	0	0.011	0	0.023	0.909
-11	0	0	0	0.013	0	0	0.087	0.013	0.025	0.013	0.138	0.7	0.013
-10	0	0	0.013	0	0	0	0.037	0	0.125	0.013	0.725	0.075	0.013
-9	0	0	0	0.012	0	0	0.024	0.024	0.145	0.699	0.048	0.048	0
-8	0	0	0	0	0.012	0.012	0.024	0.024	0.795	0.06	0.072	0	0
-7	0	0	0	0	0.009	0.009	0	0.908	0.018	0.009	0	0	0.046
-6	0	0	0	0.011	0	0	0.885	0.011	0.034	0.011	0.011	0.034	0
-5	0	0	0	0	0.035	0.814	0	0.128	0	0	0	0	0.023
-4	0	0	0	0.036	0.881	0.036	0.012	0.012	0	0.024	0	0	0
-3	0	0	0.016	0.841	0.095	0.016	0.016	0	0.016	0	0	0	0
-2	0	0.012	0.965	0	0.012	0	0	0.012	0	0	0	0	0
-1	0	0.976	0.012	0	0	0	0	0	0	0	0	0.012	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0.952	0.024	0.024	0	0	0	0	0	0	0	0
3	0	0	0	0.97	0.03	0	0	0	0	0	0	0	0
4	0	0	0	0	0.966	0	0	0	0	0.023	0	0.011	0
5	0	0	0	0	0.012	0.942	0	0.035	0	0	0.012	0	0
6	0	0	0	0	0	0	0.931	0.011	0	0	0.034	0.023	0
7	0	0	0	0	0	0.056	0	0.926	0.009	0	0	0	0.009
8	0	0	0	0.024	0	0	0.059	0.012	0.776	0.071	0.059	0	0
9	0	0	0	0	0.012	0	0	0.012	0.035	0.894	0.024	0.024	0
10	0	0	0	0.012	0	0.012	0.024	0	0.122	0.049	0.707	0.073	0
11	0	0	0	0	0	0	0.102	0.011	0.023	0.011	0.08	0.773	0
12	0.011	0	0	0	0	0	0	0.011	0	0	0	0	0.977

Table A.5. Confusion Matrix (all trials) for PN-FH

PNFH	0	1	2	3	4	5	6	7	8	9	10	11	12
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.96	0.029	0.006	0	0	0	0	0	0	0.006	0	0
2	0	0.018	0.939	0	0.012	0.006	0	0.006	0	0	0.012	0.006	0
3	0	0	0.008	0.953	0.031	0	0	0	0.008	0	0	0	0
4	0	0.006	0.012	0.035	0.871	0.024	0.006	0	0.018	0.029	0	0	0
5	0	0.006	0	0	0.012	0.918	0	0.053	0	0.006	0	0	0.006
6	0	0.006	0	0.006	0	0.012	0.844	0.029	0.017	0.006	0.029	0.052	0
7	0	0	0	0	0.014	0.06	0	0.894	0.005	0.005	0	0.005	0.019
8	0	0	0	0.012	0.006	0.018	0.042	0.018	0.777	0.042	0.066	0.012	0.006
9	0	0	0	0.006	0	0	0.006	0.018	0.072	0.837	0.036	0.006	0.018
10	0	0	0	0	0	0	0.012	0.006	0.159	0.082	0.706	0.035	0
11	0	0.006	0	0	0.006	0	0.06	0	0.024	0.024	0.089	0.792	0
12	0.006	0	0	0	0	0.006	0	0.011	0.006	0	0.011	0.011	0.949

Table A.6. Confusion Matrix by direction (of semitones) for PN-FH

PNFH	0	1	2	3	4	5	6	7	8	9	10	11	12
-12	0	0	0	0	0	0.011	0	0.023	0	0	0.023	0.023	0.92
-11	0	0.012	0	0	0.012	0	0.094	0	0.035	0.024	0.082	0.741	0
-10	0	0	0	0	0	0	0.023	0.012	0.163	0.081	0.663	0.058	0
-9	0	0	0	0.012	0	0	0.012	0.037	0.085	0.768	0.061	0	0.024
-8	0	0	0	0.012	0.012	0	0.048	0.012	0.771	0.048	0.072	0.024	0
-7	0	0	0	0	0.009	0.047	0	0.907	0	0	0	0.009	0.028
-6	0	0	0	0.011	0	0.011	0.875	0.023	0.023	0.011	0.023	0.023	0
-5	0	0.012	0	0	0.012	0.893	0	0.071	0	0.012	0	0	0
-4	0	0.012	0	0.06	0.855	0.036	0.012	0	0.012	0.012	0	0	0
-3	0	0	0.016	0.938	0.031	0	0	0	0.016	0	0	0	0
-2	0	0.024	0.94	0	0.012	0	0	0	0	0	0.012	0.012	0
-1	0	0.943	0.034	0.011	0	0	0	0	0	0	0.011	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.977	0.023	0	0	0	0	0	0	0	0	0	0
2	0	0.012	0.938	0	0.012	0.012	0	0.012	0	0	0.012	0	0
3	0	0	0	0.969	0.031	0	0	0	0	0	0	0	0
4	0	0	0.023	0.011	0.885	0.011	0	0	0.023	0.046	0	0	0
5	0	0	0	0	0.012	0.942	0	0.035	0	0	0	0	0.012
6	0	0.012	0	0	0	0.012	0.812	0.035	0.012	0	0.035	0.082	0
7	0	0	0	0	0.018	0.073	0	0.881	0.009	0.009	0	0	0.009
8	0	0	0	0.012	0	0.036	0.036	0.024	0.783	0.036	0.06	0	0.012
9	0	0	0	0	0	0	0	0	0.06	0.905	0.012	0.012	0.012
10	0	0	0	0	0	0	0	0	0.155	0.083	0.75	0.012	0
11	0	0	0	0	0	0	0.024	0	0.012	0.024	0.096	0.843	0
12	0.011	0	0	0	0	0	0	0	0.011	0	0	0	0.977

Table A.7 . Confusion Matrix (all trials) for MB-VN

MBVN	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0.977	0	0	0	0	0	0	0	0	0	0	0	0.023
1	0.012	0.936	0.023	0.017	0	0	0	0	0	0	0	0.012	0
2	0.012	0.012	0.916	0.018	0.018	0	0	0	0.006	0	0.006	0.012	0
3	0	0	0.016	0.93	0.039	0	0.016	0	0	0	0	0	0
4	0.006	0	0	0.102	0.85	0	0	0.006	0	0.006	0.018	0	0.012
5	0	0	0.006	0	0.006	0.892	0.006	0.08	0	0.011	0	0	0
6	0	0.018	0	0.012	0.018	0.018	0.847	0.012	0.024	0.006	0.012	0.035	0
7	0	0	0	0	0.019	0.046	0	0.921	0.005	0	0	0	0.009
8	0.006	0.012	0	0.006	0.018	0.012	0.047	0.029	0.731	0.07	0.035	0.012	0.023
9	0	0	0	0.006	0	0	0.006	0.018	0.09	0.825	0.03	0.024	0
10	0	0	0	0.006	0.006	0.012	0.024	0.006	0.139	0.036	0.697	0.048	0.024
11	0	0	0	0	0.006	0.006	0.041	0.012	0.053	0.024	0.101	0.751	0.006
12	0.017	0	0	0	0	0	0	0.017	0	0.006	0.006	0.006	0.948

Table A.8. Confusion Matrix by direction (of semitones) for MB-VN

MBVN	0	1	2	3	4	5	6	7	8	9	10	11	12
-12	0.023	0	0	0	0	0	0	0.011	0	0	0.011	0	0.954
-11	0	0	0	0	0.012	0.012	0.024	0.024	0.06	0.024	0.133	0.711	0
-10	0	0	0	0.013	0	0	0.037	0.013	0.125	0.037	0.738	0.037	0
-9	0	0	0	0.012	0	0	0.012	0.024	0.107	0.75	0.06	0.036	0
-8	0.012	0.023	0	0.012	0.012	0	0.047	0.023	0.721	0.081	0.047	0.012	0.012
-7	0	0	0	0	0.018	0.037	0	0.936	0	0	0	0	0.009
-6	0	0.034	0	0.011	0.023	0.023	0.793	0.023	0.034	0.011	0.011	0.034	0
-5	0	0	0	0	0	0.886	0.011	0.102	0	0	0	0	0
-4	0.012	0	0	0.159	0.805	0	0	0.012	0	0	0.012	0	0
-3	0	0	0.031	0.891	0.062	0	0.016	0	0	0	0	0	0
-2	0.012	0.012	0.94	0	0.012	0	0	0	0.012	0	0	0.012	0
-1	0.023	0.919	0.023	0.012	0	0	0	0	0	0	0	0.023	0
0	0.977	0	0	0	0	0	0	0	0	0	0	0	0.023
1	0	0.953	0.023	0.023	0	0	0	0	0	0	0	0	0
2	0.012	0.012	0.89	0.037	0.024	0	0	0	0	0	0.012	0.012	0
3	0	0	0	0.969	0.015	0	0.015	0	0	0	0	0	0
4	0	0	0	0.047	0.894	0	0	0	0	0.012	0.024	0	0.024
5	0	0	0.011	0	0.011	0.898	0	0.057	0	0.023	0	0	0
6	0	0	0	0.012	0.012	0.012	0.904	0	0.012	0	0.012	0.036	0
7	0	0	0	0	0.019	0.056	0	0.907	0.009	0	0	0	0.009
8	0	0	0	0	0.024	0.024	0.047	0.035	0.741	0.059	0.024	0.012	0.035
9	0	0	0	0	0	0	0	0.012	0.073	0.902	0	0.012	0
10	0	0	0	0	0.012	0.024	0.012	0	0.153	0.035	0.659	0.059	0.047
11	0	0	0	0	0	0	0.058	0	0.047	0.023	0.07	0.791	0.012
12	0.011	0	0	0	0	0	0	0.023	0	0.011	0	0.011	0.943

Table A.9 . Confusion Matrix (all trials) for VN-MB

VNMB	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0.977	0.011	0	0	0	0	0	0	0	0	0	0	0.011
1	0	0.943	0.034	0.006	0	0	0	0	0	0	0	0.011	0.006
2	0	0.036	0.887	0.03	0.024	0	0.012	0	0	0	0.006	0.006	0
3	0	0.023	0	0.947	0.015	0	0	0	0.015	0	0	0	0
4	0	0	0	0.048	0.88	0.018	0	0.006	0.024	0.018	0	0	0.006
5	0	0.006	0	0.006	0.018	0.887	0.012	0.048	0	0	0.006	0	0.018
6	0	0	0	0.012	0.006	0	0.864	0.012	0.018	0.006	0.036	0.047	0
7	0	0	0	0.005	0.014	0.055	0	0.894	0.009	0	0	0	0.023
8	0	0.006	0	0	0.018	0.012	0.03	0.024	0.774	0.079	0.037	0.012	0.006
9	0	0	0	0.006	0	0.012	0.029	0.017	0.058	0.802	0.052	0.023	0
10	0	0.006	0	0.006	0.006	0.012	0.012	0.012	0.117	0.058	0.713	0.053	0.006
11	0	0.012	0	0	0	0.006	0.03	0	0.041	0.012	0.13	0.763	0.006
12	0	0	0	0	0	0	0	0.058	0	0	0	0.012	0.93

Table A.10. Confusion Matrix by direction (of semitones) for VN-MB

VNMB	0	1	2	3	4	5	6	7	8	9	10	11	12
-12	0	0	0	0	0	0	0	0.094	0	0	0	0.024	0.882
-11	0	0.024	0	0	0	0	0.048	0	0.06	0.012	0.12	0.723	0.012
-10	0	0	0	0.012	0.012	0	0.024	0	0.129	0.035	0.706	0.071	0.012
-9	0	0	0	0.012	0	0.012	0.024	0.035	0.082	0.718	0.071	0.047	0
-8	0	0.012	0	0	0	0	0.024	0.024	0.759	0.12	0.048	0.012	0
-7	0	0	0	0	0.009	0.037	0	0.908	0.018	0	0	0	0.028
-6	0	0	0	0.024	0.012	0	0.869	0.012	0.012	0	0.048	0.024	0
-5	0	0	0	0.012	0.036	0.821	0.024	0.083	0	0	0.012	0	0.012
-4	0	0	0	0.072	0.88	0.012	0	0.012	0.012	0	0	0	0.012
-3	0	0.046	0	0.892	0.031	0	0	0	0.031	0	0	0	0
-2	0	0.037	0.89	0.049	0.024	0	0	0	0	0	0	0	0
-1	0	0.943	0.034	0.011	0	0	0	0	0	0	0	0	0.011
0	0.977	0.011	0	0	0	0	0	0	0	0	0	0	0.011
1	0	0.943	0.034	0	0	0	0	0	0	0	0	0.023	0
2	0	0.035	0.884	0.012	0.023	0	0.023	0	0	0	0.012	0.012	0
3	0	0	0	1	0	0	0	0	0	0	0	0	0
4	0	0	0	0.024	0.88	0.024	0	0	0.036	0.036	0	0	0
5	0	0.012	0	0	0	0.952	0	0.012	0	0	0	0	0.024
6	0	0	0	0	0	0	0.859	0.012	0.024	0.012	0.024	0.071	0
7	0	0	0	0.009	0.018	0.073	0	0.881	0	0	0	0	0.018
8	0	0	0	0	0.037	0.025	0.037	0.025	0.79	0.037	0.025	0.012	0.012
9	0	0	0	0	0	0.011	0.034	0	0.034	0.885	0.034	0	0
10	0	0.012	0	0	0	0.023	0	0.023	0.105	0.081	0.721	0.035	0
11	0	0	0	0	0	0.012	0.012	0	0.023	0.012	0.14	0.802	0
12	0	0	0	0	0	0	0	0.023	0	0	0	0	0.977

Table A.11. Confusion Matrix (all trials) for BC-MT

BCMT	0	1	2	3	4	5	6	7	8	9	10	11	12
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.965	0.006	0.017	0	0	0.006	0	0	0	0	0.006	0
2	0	0.024	0.929	0.03	0	0	0.012	0.006	0	0	0	0	0
3	0	0	0.015	0.969	0.008	0	0	0	0	0.008	0	0	0
4	0	0	0.012	0.041	0.912	0.023	0	0	0.012	0	0	0	0
5	0	0	0	0.006	0.041	0.912	0	0.035	0.006	0	0	0	0
6	0	0	0	0	0	0	0.905	0.024	0.012	0	0.041	0.018	0
7	0	0	0	0	0.023	0.047	0	0.911	0.005	0	0	0	0.014
8	0	0	0	0.019	0	0.019	0.049	0.031	0.765	0.068	0.031	0.019	0
9	0	0	0	0.012	0.012	0.018	0.024	0.012	0.03	0.805	0.065	0.024	0
10	0	0	0.006	0.006	0	0.006	0.025	0.012	0.136	0.068	0.698	0.043	0
11	0	0.006	0	0	0	0	0.098	0.012	0.018	0.024	0.079	0.762	0
12	0	0	0	0	0	0.006	0	0.034	0.006	0.006	0.006	0.006	0.937

Table A.12. Confusion Matrix by direction (of semitones) for BC-MT

BCMT	0	1	2	3	4	5	6	7	8	9	10	11	12
-12	0	0	0	0	0	0.012	0	0.047	0	0	0	0.012	0.93
-11	0	0.013	0	0	0	0	0.15	0.025	0.025	0.013	0.062	0.713	0
-10	0	0	0.012	0.012	0	0.012	0.024	0.024	0.134	0.073	0.646	0.061	0
-9	0	0	0	0.024	0	0.024	0.048	0.012	0.024	0.75	0.083	0.036	0
-8	0	0	0	0.026	0	0.013	0.052	0.026	0.753	0.065	0.052	0.013	0
-7	0	0	0	0	0.037	0.047	0	0.888	0.009	0	0	0	0.019
-6	0	0	0	0	0	0	0.915	0.024	0.012	0	0.037	0.012	0
-5	0	0	0	0.012	0.047	0.929	0	0.012	0	0	0	0	0
-4	0	0	0.012	0.059	0.871	0.047	0	0	0.012	0	0	0	0
-3	0	0	0.031	0.954	0	0	0	0	0	0.015	0	0	0
-2	0	0.024	0.905	0.048	0	0	0.012	0.012	0	0	0	0	0
-1	0	0.954	0.011	0.011	0	0	0.011	0	0	0	0	0.011	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.977	0	0.023	0	0	0	0	0	0	0	0	0
2	0	0.024	0.952	0.012	0	0	0.012	0	0	0	0	0	0
3	0	0	0	0.985	0.015	0	0	0	0	0	0	0	0
4	0	0	0.012	0.023	0.953	0	0	0	0.012	0	0	0	0
5	0	0	0	0	0.035	0.895	0	0.058	0.012	0	0	0	0
6	0	0	0	0	0	0	0.897	0.023	0.011	0	0.046	0.023	0
7	0	0	0	0	0.009	0.047	0	0.934	0	0	0	0	0.009
8	0	0	0	0.012	0	0.024	0.047	0.035	0.776	0.071	0.012	0.024	0
9	0	0	0	0	0.024	0.012	0	0.012	0.035	0.859	0.047	0.012	0
10	0	0	0	0	0	0	0.025	0	0.138	0.062	0.75	0.025	0
11	0	0	0	0	0	0	0.048	0	0.012	0.036	0.095	0.81	0
12	0	0	0	0	0	0	0	0.023	0.011	0.011	0.011	0	0.943

Table A.13. Confusion Matrix (all trials) for MT-BC

MTBC	0	1	2	3	4	5	6	7	8	9	10	11	12
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.965	0.023	0.006	0.006	0	0	0	0	0	0	0	0
2	0.006	0.035	0.901	0.035	0.006	0.017	0	0	0	0	0	0	0
3	0	0.008	0	0.961	0.016	0.008	0	0.008	0	0	0	0	0
4	0	0	0	0.024	0.941	0.012	0.006	0	0	0.018	0	0	0
5	0	0	0	0.006	0.012	0.906	0	0.065	0	0.006	0	0.006	0
6	0	0	0	0.006	0	0	0.884	0.023	0.006	0	0.046	0.035	0
7	0	0	0	0	0.014	0.046	0	0.899	0.009	0.009	0.005	0	0.018
8	0	0	0	0	0.018	0.024	0.024	0.036	0.79	0.06	0.036	0.006	0.006
9	0	0	0	0	0.006	0.006	0.012	0.006	0.083	0.845	0.018	0.012	0.012
10	0	0	0	0	0.006	0.012	0.012	0.006	0.179	0.043	0.698	0.043	0
11	0	0	0.006	0.012	0	0	0.064	0	0.064	0.006	0.139	0.699	0.012
12	0.012	0	0	0	0	0	0	0.03	0.006	0.006	0	0	0.947

Table A.14. Confusion Matrix by direction (of semitones) for MT-BC

MTBC	0	1	2	3	4	5	6	7	8	9	10	11	12
-12	0	0	0	0	0	0	0	0.037	0	0.012	0	0	0.951
-11	0	0	0.012	0.023	0	0	0.058	0	0.081	0	0.174	0.64	0.012
-10	0	0	0	0	0	0.025	0.013	0	0.175	0.05	0.675	0.062	0
-9	0	0	0	0	0.012	0	0.025	0	0.099	0.79	0.037	0.012	0.025
-8	0	0	0	0	0	0.012	0.012	0.024	0.793	0.073	0.073	0.012	0
-7	0	0	0	0	0.009	0.037	0	0.898	0.009	0.009	0	0	0.037
-6	0	0	0	0	0	0	0.862	0.034	0.011	0	0.057	0.034	0
-5	0	0	0	0.012	0	0.892	0	0.096	0	0	0	0	0
-4	0	0	0	0.049	0.915	0.012	0	0	0	0.024	0	0	0
-3	0	0.016	0	0.938	0.031	0	0	0.016	0	0	0	0	0
-2	0	0.024	0.906	0.047	0	0.024	0	0	0	0	0	0	0
-1	0	0.965	0.023	0	0.012	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0.966	0.023	0.011	0	0	0	0	0	0	0	0	0
2	0.011	0.046	0.897	0.023	0.011	0.011	0	0	0	0	0	0	0
3	0	0	0	0.985	0	0.015	0	0	0	0	0	0	0
4	0	0	0	0	0.966	0.011	0.011	0	0	0.011	0	0	0
5	0	0	0	0	0.023	0.92	0	0.034	0	0.011	0	0.011	0
6	0	0	0	0.012	0	0	0.907	0.012	0	0	0.035	0.035	0
7	0	0	0	0	0.018	0.055	0	0.899	0.009	0.009	0.009	0	0
8	0	0	0	0	0.035	0.035	0.035	0.047	0.788	0.047	0	0	0.012
9	0	0	0	0	0	0.011	0	0.011	0.069	0.897	0	0.011	0
10	0	0	0	0	0.012	0	0.012	0.183	0.037	0.72	0.024	0	0
11	0	0	0	0	0	0	0.069	0	0.046	0.011	0.103	0.759	0.011
12	0.023	0	0	0	0	0	0	0.023	0.011	0	0	0	0.943

Table B.1. Response Times for Unisons by Timbre Pair (LogRT left, RT in seconds, right)

TimPair	Mean (LogRT)	Std. Error	95% Confidence Interval		Mean (seconds)
			Lower Bound	Upper Bound	
PN-PN	.200	.033	.131	.270	1.222
FH-PN	.178	.041	.092	.263	1.194
PN-FH	.166	.042	.077	.254	1.179
MB-VN	.238	.042	.150	.325	1.268
VN-MB	.196	.040	.113	.280	1.216
BC-MT	.248	.053	.138	.357	1.281
MT-BC	.192	.058	.071	.313	1.211

References

- Allen, E. J., & Oxenham, A. J. (2014). Symmetric interactions and interference between pitch and timbre. *The Journal of the Acoustical Society of America*, 135, 1371-1379.
- Andreou, L-V., Kashino, M., & Chait, M. (2011). The role of temporal regularity in auditory segregation. *Hearing Research*, 280, 228-235.
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3(2), 12-28.
- Bailey, K. (1991). *The twelve-note music of Anton Webern*. Cambridge: Cambridge University Press.
- Beal, A. (1985). The skill of recognizing musical structures. *Memory & Cognition*, 13, 405-415.
- Bendixen, A., Denham, S. L., Gyimesi, K., & Winkler, I. (2010). Regular patterns stabilize auditory streams. *The Journal of the Acoustical Society of America*, 128(6), 3658-3666.
- Bigand, E., & Tillmann, B. (2005). Effect of context on the perception of pitch structures. In Plack, C. J., Oxenham, A. J., Fay, R. R., Popper, A. N, (Eds), *Pitch Neural Coding and Perception*, (pp. 306-351). New York, NY: Springer Science+Business Media, Inc.
- Borchert, E. M. O., Micheyl, C., & Oxenham, A. (2011). Perceptual grouping affects pitch judgments across time and frequency. *Journal of Experimental Psychology: Human perception and performance*, 37(1), 257-269.
- Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: The MIT press.
- Bregman, A. S., Ahad, A. P., Crum, P. A., & O'Reilly, J. (2000). Effects of time intervals and tone durations on auditory stream segregation. *Perception and Psychophysics* 62(3), 626-636.
- Campbell, M. "Timbre (i)." *Grove Music Online*. *Oxford Music Online*. Oxford University Press. Web. 3 July. 2015.
- <<http://www.oxfordmusiconline.com/subscriber/article/egrove/music/27973>>.
- Caplin, William (1998). *Classical form*. New York, NY: Cambridge University Press.
- Carter, E. (1985). *Esprit rude/Esprit doux*. New York, NY: Hendon Music, Boosey & Hawkes.

- _____. (1996). *Rigmarole*. New York, NY: Hendon Music, Boosey & Hawkes.
- _____. *Elliott Carter: Chamber music for winds*. Ensemble Contrasts. COP, 1998. CD.
- _____. *Elliott Carter*. Nouvel Ensemble Moderne, Lorraine Vaillancourt. ATMA Classique, 2000. CD
- _____. (2002). *Au Quai*. Milwaukee, WI: Hendon Music, Boosey & Hawkes.
- Caruso, V., & Balaban, E. (2014). Pitch and timbre interfere when both are parametrically varied. *PLOS ONE*, 9. doi: 10.1371/journal.pone.0087065.
- Coulembier, K. (2009). Complexity as compound simplicity: Elliott Carter's "Esprit Rude/Esprit Doux." *Revue belge de Musicologie*, 63, 147-161.
- 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. doi: 10.3389/fnsys.2013.00127.
- Cramer, A. (2002). Schoenberg's Klangfarbenmelodie: A principle of early atonality. *Music Theory Spectrum*, 24, 1-34. doi: 10.1525/mts.2002.24.1.1.
- Dahlhaus, Carl. (1987). *Schoenberg and the new music*. Trans. Derrick Puffett & Alfred Clayton. Cambridge, UK: Cambridge University Press.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, M.D: Erlbaum.
- Handel, S., Weaver, M. S., & Lawson, G. (1983). Effect of rhythmic grouping on stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, 9(4), 637-651.
- Hopkins, N., & Link, J.F. (2002). *Elliott Carter harmony book*. New York, NY: Carl Fischer, LLC.
- Iverson, J. (2009). *Historical memory and Gyorgi Ligeti's sound mass music 1958-1968*. Unpublished Doctoral Dissertation, University of Texas at Austin, USA.
- Jones, D. (2010). A WEIRD view of human nature skews psychologists' studies. *Science*, 328, 1627.
- Killam, R. N., Lorton, P. V, Jr., & Schubert, E. D. (1975). Interval recognition: identification of harmonic and melodic intervals. *Journal of Music Theory*, 19, 212-234.
- Krumhansl, C., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 739-751.

- Marozeau, J., de Cheveigné, A., McAdams, S., & Windsberg, S. (2003). The dependency of timbre on fundamental frequency. *The Journal of the Acoustical Society of America*, 114, 2946-2957.
- Martin, F. N., & Champlin, C. A. (2000). Reconsidering the limits of normal hearing. *Journal of the American Academy of Audiology*, 11(2), 64–66.
- Marvin, E. W., & Brinkman, A. R. (2000). The effect of key color and timbre on absolute pitch recognition in musical contexts. *Music Perception: An Interdisciplinary Journal*, 18, 111-137.
- Mathews, P., ed. (2006). *Orchestration: An anthology of writings*. New York, NY: Routledge, Taylor & Francis Group.
- McAdams, S. (2013). Musical timbre perception. In Diana Deutsch (Ed.), *The Psychology of Music*, 3rd Edition, (pp. 35-67). San Diego: Elsevier.
- McAdams, S., & Bregman, A. S. (1979). Hearing Musical Streams. *Computer Music Journal*, 3(4), 26-61.
- McAdams, S., Winsberg, S., Donnodieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58, 177-192.
- Melara, R. D., & Marks, L. (1990a). Hard and soft interacting dimensions: Differential effects of dual context on classification. *Perception & Psychophysics*, 47, 307-325.
- Melara, R. D., & Marks, L. (1990b). Interaction among auditory dimensions: Timbre, pitch and loudness. *Perception and Psychophysics*, 48, 169-178.
- Melara, R. D., & Marks, L. (1990c). Perceptual primacy of dimensions: Support of a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 398-414.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, 219, 36-47.
- Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *ACTA Acustica United with Acustica*, 88, 320-332.
- Oxenham, A. J. (2013). The perception of musical tones. In Diana Deutsch (Ed.), *The Psychology of Music*, 3rd Edition, (pp. 35-67). San Diego: Elsevier.

- Oxenham, A. J., Fligor, B. J., Mason, C. R., & Kidd, G. Jr. (2003). Informational masking and musical training. *The Journal of the Acoustical Society of America*, 114 (3), 1543-1549.
- Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The Timbre Toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, 130(5), 2902-2916.
- Pitt, M. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 976-986.
- Rakowski, A. (1985). The perception of musical intervals by music students. *Bulletin of the Council for Research in Music Education*, 85, 175-186.
- Rogers, W. L., & Bregman, A. (1993). An experimental evaluation of three theories of auditory stream segregation. *Perception & Psychophysics*, 53(2), 179-189.
- Russo, F., & Thompson, W. F. (2005a). The subjective size of melodic intervals over a two-octave range. *Psychonomic Bulletin & Review*, 12, 1068-1075.
- Russo, F., & Thompson, W. F. (2005b). An interval size illusion: The influence of timbre on the perceived size of melodic intervals. *Perceptions and Psychophysics*, 67, 559-58.
- Samplaski, A. (2005). Interval and interval class similarity: Results of a confusion study. *Psychomusicology*, 19, 59-74.
- Schiff, D. (1998). *The music of Elliott Carter* (2nd ed.). Ithaca, NY: Cornell University Press.
- Schoenberg, A. (1948). *Theory of harmony*. Robert D. W. Adams, trans. New York, NY: Philosophical library.
- Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *The Journal of the Acoustical Society of America*, 89, 2404-2410.
- Silbert, N. H., Townsend, J. T., & Lentz, J. J. (2009). Independence and separability in the perception of complex nonspeech sounds. *Attention, Perception, & Psychophysics*, 71 (8), 1900-1915. doi:10.3758/APP.71.8.1900.
- Singh, P. G., & Hirsh, I. J. (1992). Influence of spectral locus and F0 changes on the pitch and timbre of complex tones. *The Journal of the Acoustical Society of America*, 92(5), 2650-2661.
- Smith, B. K. (1995). PsiExp: An environment for psychoacoustic experimentation using the IRCAM musical workstation. In *Society for Music Perception and Cognition Conference '95*. Berkeley, CA: University of California, Berkeley.

- Snyder, J. S., & Weintraub, D. M. (2011). Pattern specificity in the effect of prior Δf on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1649-1656.
- Spinner, L. (1955). Analysis of a period. *Die Reihe*, 2, 46-50.
- Straus, J.N. (2009). *Twelve-tone music in America*. Cambridge, UK: Cambridge University Press.
- _____. (2005). *Introduction to post-tonal theory*, 3rd ed. Upper Saddle River, NJ: Pearson Prentice Hall.
- Truniger, M. (1998). Elliot Carter: Esprit rude/esprit doux. *Sonus*, 19, 26-52.
- Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception*, 6, 383-396.
- Vurma, A. (2014). Timbre-induced pitch shift from the perspective of Signal Detection Theory: the impact of musical expertise, silence interval, and pitch region. *Frontiers in Psychology*, 5, 1-13. doi: 10.3389/fpsyg.2014.00044.
- Vurma, A., Raju, M., & Kuuda, A. (2010). Does timbre affect pitch?: Estimations by musicians and non-musicians. *Psychology of Music*, 39, 291-306. doi: 10.1177/03573561037602.
- Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics*, 64, 198-207.
- Webern, A. (1948). *Konzert, op. 24*. Vienna: Universal Edition.
- _____. (1960). *Quartett für geige, klarinette, tenorsaxophon und klavier. op. 22*. Vienna: Universal Edition.
- _____. *Anton Webern, complete works*. Ensemble Intercontemporain, Pierre-Laurent Aimard (piano). Cond. Pierre Boulez. Deutsche Grammophon, 2000. CD.
- Whittwall, A. (2008). *The Cambridge introduction to serialism*. Cambridge, UK: Cambridge University Press.
- Winsberg, S., & De Soete, G. (1993). A latent class approach to fitting the weighted Euclidean model CLASCAL. *Psychometrika*, 58(2), 315-330.
- Wintle, C. (1982). Analysis and performance of Webern's Concerto Op. 24/II. *Music Analysis*, 1, 73-79.
- Zarate, J. M., Ritson, C. R., & Poeppel, D. (2013). The effect of instrumental timbre on interval discrimination. *PLOS ONE*, 8(9). doi: 10.1371/journal.pone.0075410.