ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF ARTS AND SOCIAL SCIENCES

THE STRUCTURAL ELEMENTS OF TIMBRAL COHERENCE AND ITS COMPOSITIONAL APPLICATION

Ph.D. THESIS

Enis GÜMÜŞ

Department of Music

Music Programme

MAY 2020

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF ARTS AND SOCIAL SCIENCES

THE STRUCTURAL ELEMENTS OF TIMBRAL COHERENCE AND ITS COMPOSITIONAL APPLICATION

Ph.D. THESIS

Enis GÜMÜŞ
(409132013)

Department of Music

Music Programme

Thesis Advisor: Asst. Prof. Dr. Eray ALTINBÜKEN

MAY 2020

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ SOSYAL BİLİMLER ENSTİTÜSÜ

TINISAL BÜTÜNLÜĞÜN YAPISAL ÖĞELERİ VE KOMPOZİSYONEL UYGULAMASI

DOKTORA TEZİ

Enis GÜMÜŞ
(409132013)

Müzik Anabilim Dalı

Müzik Programı

Tez Danışmanı: Dr. Öğr. Üyesi Eray ALTINBÜKEN

MAYIS 2020

Enis GÜMÜŞ, a Ph.D. student of ITU Graduate School of Arts and Social Sciences student ID 409132013, successfully defended the dissertation entitled "THE STRUCTURAL ELEMENTS OF TIMBRAL COHERENCE AND ITS COMPOSITIONAL APPLICATION", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Asst. Prof. Dr. Eray ALTINBÜKEN** Istanbul Technical University	.....
Jury Members :	**Prof. Dr. Can KARADOĞAN** Istanbul Technical University	.....
	Prof. Dr. Tolga TÜZÜN Istanbul Bilgi University	.....
	Assoc. Prof. Dr. Jerfi AJİ Istanbul Technical University	.....
	Assoc. Prof. Ali Ahmet ALTINEL Mimar Sinan Fine Arts University	.....

Date of Submission : 30 April 2020
Date of Defense : 25 May 2020

Dünya'nın bir yerinde, milyarlarca yıldır bekleyen
ve bizden sonra da milyarlarca yıl orada olacak bir kaya parçasına.

FOREWORD

First of all, I would like to thank my advisor Asst. Prof. Dr. Eray ALTINBÜKEN, thesis committee members Prof. Dr. Tolga TÜZÜN and Prof. Dr. Can KARADOĞAN, and jury members Assoc. Prof. Dr. Jerfi AJİ and Assoc. Prof. Ali Ahmet ALTINEL for their support and guidance throughout the development and completion of this thesis. I would also like to express my gratitude to Assoc. Prof. Dr. Robert REIGLE and Dr. Reuben DE LAUTOUR for their help and support during the initial stages of my work.

I would like to thank Ali ÖZDOLAP for guiding me in my discovery of the essence of all music. I would also like to thank René UIJLENHOET, Robin DE RAAFF, Luc BREWAEYS (R.I.P.), Jan Bas BOLLEN and Assoc. Prof. Dr. Koray SAZLI from my undergraduate and master's education years, for teaching me how to embrace my profession as it is, and giving me the courage to pursue a life in it.

I am thankful to my dear friends and colleagues for their support and friendship; with special thanks to Dr. Sair Sinan KESTELLİ and Şehnaz İpek GÖRGÜN for their companionship and solidarity during my doctoral studies; to Dilara TURAN, Tunca OLCAYTO and Aslı KOBANER for enduring our tough debates of all kinds of topics; to Baran DÜZCE, Nergiz ERCAN, Dr. Kamil Korhan NAZLIBEN, Dr. Haki PAMUK, Dr. Serhan SADIKOĞLU, and Sam David WAMPER for being wonderful friends in the days of darkness; and to all fellow travelers who share the burdens of this one-way journey.

I am grateful to my father, İsa GÜMÜŞ and my brother, Murat GÜMÜŞ, for believing in me, regardless of whatever I am doing.

I am grateful to my mother, my first music teacher, Ayşe GÜMÜŞ, whom I miss a lot, for the life and love she has given me.

Lastly, I would like to thank my housemates L. Sıfır and R. Buçuk for keeping me sane (especially during these 'quarantine days'); and the great, dark and endless void, for reminding me of its conceptual existence that flows into this supposedly material world, drags everything out of their places, and carries on forever. Without the inspiration and warnings coming from it, this work would not be possible.

May 2020 Enis GÜMÜŞ

TABLE OF CONTENTS

Page

ABBREVIATIONS

ADSR	**:** Attack, Decay, Sustain, Release
ca.	**:** Circa
gest.	**:** Gesture
IRCAM	**:** Institut de Recherche et Coordination Acoustique/Musique
m., mm.	**:** Measure, measures
MDS	**:** Multi dimensional scale
no.	**:** Number
op.	**:** Opus
PAT	**:** Perceptual Attack Time
q	**:** Quarter note (crotchet)
reg.	**:** Region

SYMBOLS

F0	**:** Fundamental Frequency
k	**:** Thousand
kHZ	**:** Kilohertz
ms	**:** Milisecond(s)
s	**:** Second(s)

LIST OF TABLES

Page

LIST OF FIGURES

Page

THE STRUCTURAL ELEMENTS OF TIMBRAL COHERENCE AND ITS COMPOSITIONAL APPLICATION

SUMMARY

This thesis aims to define 'timbral coherence' in its structural elements and develop an approach for formal analysis. Comparative analysis of cognitive studies on timbre, and the philosophy of musical gesture are the main sources to achieve this goal. Analyses of these studies are used to present a proposal for building new analysis methods.

Chapter 1 presents a brief introduction of timbre, its historical and contemporary role in music, main definitions and topics concerning the nature of timbre. An initial definition of timbral coherence is proposed to determine a path and approach for following chapters.

Cognition of timbre is examined through noteworthy studies in Chapter 2. First of all, an overview of notion timbral continuum, vertical and horizontal organization in composition is made. Subsequently, relation of timbre and horizontal structures is explained starting from the most basic components: pitch, intervals, and melody. Most importantly, relativity, memory and its grouping mechanisms are examined, which can be seen as the basis for devising the proposed approach. The inferences made through these sections are used to prepare a basis for the following sections where categorization of timbral levels and types are presented.

Chapter 3 investigates musical gesture, its types and subtypes from a compositional perspective, in search of possible hierarchical relations on different levels. Examination of these relations starts from analogies and direct relations between physical and mental gestures, both in language and music. Then, a categorization of performer's gestures in instrumental practices is presented. An explanation of more complex structures and larger dimensions through possible gestural networks is given, in order to establish an idea of a higher level of gestural significance. Consequently, a consideration of compositional gesture and the factors in its conveyance and transformation is made through different stages of the composition, performance, and its reception.

In Chapter 4, an approach for three-layer analysis of formal structure is proposed by utilizing the main ideas inferred about musical gesture from the previous section in combination with timbre and memory. The concepts that are essential to the proposed approach are examined in each sub-section. Timbral object and its analysis in timbral networks are defined, broadening the initial deductions made in Chapter 2. This is followed by gestural network analysis where gesture is put into context with timbral issues. Two following sections explain the notion of markedness and contextual considerations for both analyses of timbre and gestural networks. This chapter is concluded with a fictive case analysis, using the initial description of the method.

Finally, Chapter 5 puts the theoretical information into practice by analyzing two short works, Anton Webern's Op. 10, No. 1 *(Fünf Stücke für Orchester)* and George Crumb's "Threnody I: Night of the Electric Insects", a movement from his "Black Angels" for electric string quartet. Provisional analyses of these works provide a preview of what a complete method would be and what is to be done for achieving it.

This thesis is concluded with general consideration of the covered topics and possible further work.

TINISAL BÜTÜNLÜĞÜN YAPISAL ÖĞELERİ VE KOMPOZİSYONEL UYGULAMASI

ÖZET

Bu tez, 'tınısal bütünlüğü' yapısal öğeleri içerisinde tanımlamayı ve biçimsel analiz için bir yaklaşım geliştirmeyi amaçlamaktadır. Tını üzerine bilişsel çalışmaların karşılaştırmalı analizi ve müzikal jestin felsefesi bu hedefe ulaşmak için ana kaynaklardır. Bu çalışmalardan öğrenilenler, yeni analiz yöntemleri oluşturulmasına yönelik bir öneri sunmak için kullanılmıştır. Burada bahsedilen tını ve jest temelli yaklaşımda jest, aslen var olan yerleşik analiz yaklaşımlarının genelleştirilmiş, farklı doku ve stillere uygulanabilir hale getirilmiş bir yorumudur. Tını ise daha merkezi durumdadır. Tezin temel hareket noktası, tınının günümüz müziklerinde türler-üstü bir bütünleştirici algılama öğesi olduğu ve bu noktaya tarihsel bir süreç içerisinde ulaşıldığı fikridir.

Bu ana fikir doğrultusunda, Bölüm 1, tınının müzikte bir kavram olarak kısa bir tanıtımını, tarihsel ve çağdaş rolünü, tınının doğasıyla ilgili ana tanımları ve başlıkları sunmaktadır. Sonraki bölümler için bir yol ve yaklaşım belirlemek amacıyla, tınısal bütünlüğün bir ön tanımı şu şekilde önerilmiştir: Tını bütünlüğü, bir müzik eserini dikey (anlık), yatay (zamansal) ve bu ikisinin sürekliliğini, veya tınısal veri sürekliliği oluşturabilecek ve böylelikle söz konusu müziğin algısını daha büyük ölçekte etkileyebilecek boyutları karakterize eden bir nitelik, kriter ya da parametredir. Yani bu tanım, tınısal bütünlüğün çeşitli nitelikler ve boyutlar üzerinden biçimsel algıyı şekillendirmesi fikri üzerine kuruludur. Bu düşünüşün tarihsel kökenlerini anlamak için, bu bölümün üçüncü alt başlığı tını kavramının kompozisyon tarihindeki yeri üzerine eğilmektedir. Özellikle Barok Dönem müziğindeki gelişmeleri ve çeşitliliği göz önüne alarak, hem armoni anlayışındaki yansımalara hem de çalgıların gelişimine dikkat çekilmiştir. Devamında, 20. yüzyılın ilk yarısından Arnold Schoenberg'in düşüncelerine, ikinci yarısından ise Spektral Müzik kavramına dikkat çekilmiştir.

Tını bilişi, Bölüm 2'de, kayda değer bilimsel çalışmalar üzerinden incelenmiştir. İlk olarak, tınısal süreklilikliğin, kompozisyonda dikey ve yatay organizasyon kavramların genel değerlendirmesi yapılmıştır. Bu değerlendirme, konuyla ilgili daha önceki çalışmalara atıflar ve önemli noktaları açıklamaya yönelik çeşitli kayda değer örneklerle desteklenmiştir. Bu bölümlerdeki temel amaç, öncelikle tınısal sürekliliğin fiziksel ve algısal taraflarının ne şekilde ayrı kavramlar oldukları ve nasıl değerlendirilebilecekleri konularını ortaya koymak; akabinde de, uygun müzikal örneklerle bu fikirlerin uygulamaları üzerinde nasıl analizler yapılabileceğine dair bir tartışma başlatmaktır. Bu bölümdeki örnekler, önemli spektralist besteciler Gérard Grisey ve Tristan Murail'dan alınmıştır.

İkinci Bölüm'ün üçüncü alt başlığında, tını ve yatay yapıların ilişkisi en temel bileşenlerden başlayarak açıklanmaktadır: perde, aralıklar ve melodi. İlk olarak, tınının perde ve dinamiklerle ilişkisi incelenmiştir. Bu konuda yine Schoenberg'in ses

-rengi (*klangfarbe*) kavramına bakışı, tını ve perdenin fonksiyonel ilişkisine vurgu yaparak kullanılmıştır. Bu fikri daha anlaşılır hale getirmek için, çalgılarda değişik ses bölgelerindeki özellikleri ve kullanılan dinamiklere bağlı değişikliklerin tınısal yansımaları örneklenmiştir. Burada yine kompozisyonel bir uygulama örneği olarak, George Enescu'nun *"Carillon Nocturne"* adlı eseri, çağdaş anlamdaki tınısal yaklaşımın bir öncülü olarak kısaca incelenmiştir. Bunu takip eden, 'melodik aralıklar ve entonasyon' başlıklı bölümde, bilişsel alandaki bazı çalışmalar karşılaştırılmış, özellikle spektral özelliklerin, perdelerin ve bir bütün olarak melodik karakterin algısını nasıl etkilediği sorusuna cevap aranmıştır.

'Melodi' başlığı altında takip eden diğer bölümlerde, önerilen yaklaşımın oluşturulması için en temel konular olan görelilik, hafıza ve gruplama mekanizmaları incelenmiştir. Burada özellikle 'görelilik' kavramı öne çıkmaktadır; çünkü tını algısının görelilik açısından zayıf, mutlak özelliklerinin ise baskın olduğu fikri, yöntemsel yaklaşımları temelinden etkileyebilecek derecede önem taşımaktadır. Tınının algısal mutlaklığını destekleyen deneysel çalışmalardan bahsedilmiş ve bunların olası yansımaları tartışılmıştır. Bir sonraki alt başlıkta ise gruplama mekanizmaları, *Gestalt* prensipleriyle bağlantılı şekilde anlatılmıştır. Müzik bilişi dahilindeki deneysel çalışmalardan ortaya çıkan sonuçlar ve bu sonuçlara dayanan temel prensipler örneklenmiştir. Bu bölümdeki son alt başlıkta ise beklenti ve hafıza konuları incelenmiş, hem kısa sürelerde hem de zamana yayılan süreçlerde tını algısı üzerine çalışmalar incelenmiştir. Beklenti ve kültürel eğilimler arasındaki ilişki ve bunların müzikal yapılardaki algıyı nasıl şekillendirdiği üzerine temel bir anlayış şekillendirilmiştir.

Tını bilişinin ele alındığı İkinci Bölüm'ün son kısmı, bu konuda elde edilen çıkarımların ışığında bir yeniden değerlendirme yapmakta ve yöntem geliştirme yaklaşımında ilk basamak olacak tınısal bir kategorizasyon sunmaktadır. Lerdahl'ın 1987 tarihli "Tınısal Hiyerarşiler" (*Timbral Hierarchies*) başlıklı makalesinden yola çıkarak bir sağlama yapılmakta, bu süreçte edinilen bilimsel bilgi ve yaklaşımların bizlere kazandırdıkları düşünülerek, güncellememiz gereken görüşler sorgulanmaktadır. Bölüm 2'nin bu kısmı, 'düzeyler ve tipler' bazında iki kategorizasyon önerisi sunar. Bunlar, düzeyler bazında mikro-tını (*microtimbre*), makro-tını (*macrotimbre*), toplam-tını (*overall timbre*) ve formel tını (*formal timbre*); tipler bazında (verili sınıflar üzerinden) ise kaynak kimliği, ses aralığı bölgesi (*registral region*), spektral bölge (*spectral region*), tınısal zarf (*timbral envolope*) ve modülasyonlar şeklindedir. Buradaki düzeyler, 'birinci' ve 'ikinci' olarak ikili bir hiyerarşi çerçevesinde toparlanmıştır. Bu alt başlıklar boyunca yapılan çıkarımlar, tınısal düzeylerin ve tını tiplerinin kategorizasyonunun işlevselleştirildiği Dördüncü Bölüm için bir temel hazırlamaktadır.

Bölüm 3'te, müzikal jest, türleri ve alt-türleri, farklı düzeylerdeki olası hiyerarşik ilişkilerin arayışı içerisinde, kompozisyonel açıdan incelenmektedir. Bu ilişkilerin incelenmesi, hem dil hem de müzikte, fiziksel ve zihinsel jestler arasındaki analojilerden ve doğrudan ilişkilerden başlamaktadır; çünkü dilin oluşumu ve gelişimindeki yardımcı jestlerin, müzikal jestler ile karşılaştırılması ve aralarındaki bağlantıların sorgulanması, jesti bir bütün olarak anlamak açısından önem arz eden bir konudur.

Jest bölümünün ikinci başlığında, müzikal jestin tanımlanması ve müzikal jest tiplerinin hiyerarşik bir şekilde kategorize edilmesi amaçlanmaktadır. Bu doğrultuda, ilk olarak çalgısal uygulamalarda icracı jestlerinin bir kategorizasyonu sunulmuştur.

İcracı jestleri hem Delalande'ın kategorizasyonu ile 'efektif', 'yardımcı' ve 'figüratif' jestler olarak ele alınmakta, çalgısal jestler ise Cadoz'un 'uyarım', 'modifikasyon' ve 'seçim' jestleri ile özetlenmektedir. Bu temel jest kategorilerine ek olarak, tını için de yapıldığı gibi, jestlerin geneli için iki düzeyli bir hiyerarşi önerilmiştir. 'Birinci' düzeyde, fiziksel jest kavramına en yakın şekilde ele alınan müzikal jest, işlevsellikleri üzerinden ayrıştırılmak üzere, 'figür' ve 'motif' kavramları ile karşılaştırılmıştır. Bu kısımdaki '(anlatımsal) jest birimi' (*expressive unit gesture*), tını hiyerarşisindeki mikro-tınıya karşılık gelen, en temel jestin isimlendirilmesidir. 'İkinci' düzeyin, daha karmaşık ve zaman eksenine oturan tanımı, ve daha sonra jestsel ağlarda incelenebilecek yüksek düzeyde bir jestsel ağırlık (*gestural significance*) fikri oluşturabilmek için, olası karmaşık yapıların ve daha geniş boyutların bir açıklamaları yapılmıştır. Tüm bu çözümlemelerin sonucu olarak, jestlerin hiyerarşik açından birinci ve ikinci düzeylerde, tiplendirme açısından da fiziksel ve zihinsel kategorileri altında değerlendirildiği genel bir şema sunulmuştur. Bir sonraki alt başlıkta ise öncelikle, kompozitörlerin jest kavramını nasıl algıladıkları ve kategorize ettiklerine örnekler üzerinden değinilmiş; bunun sonrasında da kompozisyonel jestin iletim ve dönüşümündeki faktörler; kompozisyon, performans ve alımlamanın farklı aşamalarında değerlendirilmiştir. Söz konusu aşamalarda kompozitör, icracı ve dinleyici özneler olarak; müzikal notasyon, ses ve görsel öğeler ise iletim ortamları olarak incelenmiştir.

Dördüncü Bölüm'de, tını ve bellek ile birlikte, bir önceki bölümden müzikal jest hakkında edinilen ana fikirlerden faydalanılarak, biçimsel yapının üç-katmanlı (tını-jest-hafıza) analizi için bir yaklaşım önerilmiştir. Mümkün olduğunca sağlam ve meşru temellere dayandırılmaya çalışılan bu yaklaşım, 'nesnellik' prensibinin ne şekilde ele alınabileceğini tartışarak inşa edilmeye başlanmıştır. Bu noktada, kültürel altyapının algıya etkisindeki temel yaklaşım biraz daha açılmış, esas amaç ifade edilmiştir. Öncelikle yöntemin genel bir tarifi yapılmış, ardından, yöntem inşasında gerekli olan kavramlar alt bölümlerde teker teker incelenmiş, ve bu kavramların işlevselleştirilebilmesine yönelik fikir yürütülmüştür. Tını nesnesi ve tını nesnesinin ağsal analizi, İkinci Bölüm'deki çıkarımları genişletmek yoluyla tanımlanmıştır. Burada, söz konusu kavramların nasıl ifade edilip, işlevselleştirilebilecekleriyle ilgili seçenekler sunulmuştur, tınısal düzeyler ve tipler arasındaki ilişkiler netleştirilmiştir.

Dördüncü Bölüm'ün üçüncü alt başlığında, jestin tınısal meseleler ile bağlamsallaştırıldığı 'jestsel ağ analizi' bulunmaktadır. Burada jest ve 'tınısal jest' arasındaki bağlantılar ve jeste uygulanabilen transpozisyon ve mutasyon, son olarak da 'zamansal değerlendirme' kavramları açıklanmıştır. Sonraki kısımlarda hem tını hem de jest ağlarının analizleri için önem taşıyan 'işaretlenmiştlik' (*markedness*) ve 'bağlamsal değerlendirme' kavramları yer almaktadır. İşaretlenmişlik, aslen dilbilime ait olmasına rağmen, müzik alanında da kullanılabilen bir kavram olarak, bu çalışmaya en işlevsel araçlardan biri olarak dahil edilmiştir. 'Bağlamsal değerlendirme' aşaması ise, şekilsel analizi çok benzeyen öğelerin farklı bağlamlarda nasıl işlevlerinin ve etkilerinin değiştiği konusunu tartışarak, işaretlenmişlik kavramını tamamlamaktadır. Dördüncü Bölüm, analiz yönteminin ilk tarifine dayanan kurgusal bir vaka analizi ile sonuçlandırılmıştır. Burada analiz yöntemindeki olası geçici aşamalar doğrultusunda ortaya çıkabilecek bir şema sunulmuş ve şemanın nasıl okunabileceği açıklanmıştır.

Son olarak, Beşinci Bölüm'de, Anton Webern'in Op. 10, No. 1 (*Fünf Stücke für Orchester)* ve George Crumb'ın elektrikli yaylı dörtlüsü için bestelediği "*Threnody I: Night of the Electric Insects*" *(Black Angels)* adlı eserleri analiz edilerek, bu noktaya kadar hazırlanan teorik bilgiler uygulamaya konulmuştur. Bu uygulama sonucunda

hem form şemaları oluşturulmuş hem de bir yazılı notasyon sistemi önerilmiştir. Söz konusu eserlerin bu nihai olmayan analizleri, tamamlanmış bir yöntemin ne olacağı ve bunu başarmak için ne yapılması gerektiği konusunda bir ön-izleme olarak görülebilir. Bu çaba doğrultusunda, birbirinden farklı dokularda eserler seçilmiş, fakat anlatım dilinin ortaklaştırılması hedeflenmiştir.

Tez, üzerinde çalışılan ve bu çalışmanın ötesinde ele alınabilecek muhtemel konuların genel değerlendirilmesi ile sonuçlandırılmıştır.

1. INTRODUCTION

After the impact of the Second Industrial Revolution followed by further urbanization at the beginning of the 20th century, sound world of the urban society has changed significantly. As we can see the hints of these changes in early 20th century writings and artworks, the mankind confused by the rapid development of technology and social change, responded in many ways. The long 19th century 'ending' with a *Fin de siècle* and being labelled as *La Belle Époque*[1] retrospectively; a turbulent era consisting of two world wars followed by The Cold War and The Digital Age; humanity sustained a constant mode of confusion, destruction and development at the same time. The connotations of possibilities extending in a wide range of interpretations have become the default mode for every field, regardless of their perceptions in negative or positive.

The 20th century encompassed futurists of Italy and Soviet Union, Second Viennese School, sound liberators, electronic music, minimalism, spectralism and many more movements that are related to the timbral world and its music. Only in one century we have come to the concepts of post-industrial, even post-digital soundscapes; and we were never able to hear and perceive anything as we did before. Instruments continued to develop, and music became timbrally richer. In the post-tonal era of Western Art Music, in the absence or scarcity of tonal/pitch based/traditional elements, textural density and timbre have become most important aspects for identifying contemporary music.

Today, modern urban soundscape works as a journey through 'colorful sounds' or 'noises', that continuously expose the listener to the continuous and discontinuous sets of them. These words come with many connotations depending on every individual's

[1] These time-period related terms (The long 19th century, *Fin de siècle*, *La Belle Époque*) refer to the development of related and contradicting ideas about what brought humanity to the 20th century. 'Long 19th century' refers to the period between French Revolution and the beginning of the WWI. '*Fin de siècle'* (End of century) refers to the confusion and belief that humanity is in a serious decay, degeneration; and the term '*La Belle Époque'* (Beautiful Epoch) was used retrospectively for approximately the same time period.

profile and preferences. In other words, as a result of this sound inflation, we are looking for the proper soundscapes, and as one of the main parts of it, musics which are fit for listening and gathering around coherent wholes, meanings. Perhaps, timbres are analogous to scents, creating strong connections through our memory of sound and music culture.

Many works from 20th and 21st centuries lacked the expectancy factor in a traditional, formulized way, in a sense that tonal music had before; or expectancy has become directed to parameters other than pitch related materials. Furthermore, as the techniques, mechanisms, styles of the music diversify; the elements that shape the perception of musical structures and form are directed towards more universal aspects. Here, the universal is meant to indicate the common cognitive features of human beings, rather than a cultural notion of universal. The reason behind this assumption will be explained in the following sections.

1.1 Motivation and Purpose

Being a musician of several genres of music and being educated as a musician/composer in both Turkish Music and Western Art Music traditions, I often had the experience of conflicting (or coexisting but separate) perceptions of musical structures and form; or witnessed similarities and differences in others'.

Appreciation of form is a subject of aesthetics; and aesthetics look for the 'beauty'. Beauty in nature serves a very simple function, to reproduce, to replicate genes. Similarly, a 'meme'[2] (in this case, musical structures and form) replicates itself through cultural means. Music is a complex of memes. It replicates and evolves through its structural and formal features that are resilient. So, when we analyze music, we analyze the structures that shape the form. According to Murail and Risset, it is important "to be concerned with the relation between the conception and perception" (Moscovich, 1997, p. 22). This relation inevitably must be examined in both

[2] Meme: Any cultural unit with the potential of self-replication and mutation while it spreads from person to person. First coined by Richard Dawkins in his 1976 book "The Selfish Gene" and further developed by Susan Blackmore in her 2000 book "The Meme Machine".

directions. Therefore, the study and the craft of composition must be concerned with a robust and objective[3] methods of analysis.

Purpose of this thesis is to establish a functional definition of 'timbral coherence' in order to make inferences for compositional and theoretical uses; and set underlying principles for a feasible and coherent analysis/synthesis approach. The notion of 'timbral coherence' in all possible dimensions and meanings will be attempted to be covered. In addition, in order to complete timbral references in a greater network of structures, a second notion that is crucial to its analysis (and in general, for music) will be examined in a broader sense: gesture. Gestural analysis is not essentially new, more conventional methods of analysis use gestural aspects of motifs and other structural elements in different ways. The concept of gesture represented here is specifically tailored for compositional use.

This thesis attempts to set a philosophical approach for new methods and does not directly assert a specific one. It is completely on the analysis of Western Art Music and its listener; however, relevant connections to other genres will be examined when needed.

1.2 Timbral Coherence

Timbre is quite an important aspect of our daily auditory environment since it is the main property of the sound that enables us to recognize the sources we hear. In an evolutionary perspective, this was, and still is, an essential ability for survival of our species like many other animals. Thus, educated or uneducated, every ear is equipped with timbral expertise to some extent.

François (1990) relates the 'trigger timbre', a fixed timbre, to visuality and past references; he connects the idea of timbre with our visual memory and gesture perception (p. 116). This idea can be seen as an extension of the effects of the recording technology and computers in music; and this is directly related to our perception of any music. The availability of repetition leads to memorization; and thus, timbre becomes 'frozen'. "What was interpretation is now frozen, so that the structural aspects of the work tend to disappear to become a pure ornament-that is, a timbre

[3] About the adjective 'objective', see Chapter 4.

object" (François, 1990, p. 114). The 'absoluteness' of timbre becomes a crucial notion at this point. Its strength as a memory object and its relation with gesture is essential in the search of timbral coherence; and this relation and its components will be examined in the following chapters.

The notion of timbral coherence can be firmly defined after a contextual exploration. Nonetheless, I propose a provisional definition: (In musical context) timbral coherence is a quality, criterion or a parameter that characterizes a given music in vertical (instantaneous), horizontal (temporal) and the continuum of these two; or any other possible dimension that conveys a continuum of timbral data, and influences the perception of that music in greater scale.

Before going further into the notion of timbre, two other terms should be briefly explained in order to avoid confusion about the definition of timbral coherence. One of them is 'timbral blend', which directly refers to a simultaneous auditory fusion of timbral elements in a specific context (e.g., instruments in an ensemble or any recorded sounds played simultaneously). By this definition, timbral blend can be seen as a sub-topic of timbral coherence; because, as it will be presented in the following chapters, timbral coherence is sought not only in short durations but in extended formal relations.

The second term that should be mentioned is texture. Texture is a combination of instrumental timbres, rhythms, placement of pitches, placement of gestures and all possible elements. Most importantly, unlike timbre, auditory analysis and explanation of texture is relatively simple. On the other hand, timbre is audible, but not easily describable. Textural and timbral structures can be grouped in very similar sets, but their implications are separate.

1.3 A Brief History of Timbre in Composition

The role of timbre in the Western Art Music tradition has never been primary during most of its history. It was a result that is caused by other structures in focus, or an ancillary function, an ornamentation. As François (1990) states, it had to "be present without applying too much weight" (p. 115). The timbral results of orchestration were there only to help the pitch content, harmony to be perceived in a certain, clear way. This has changed throughout the centuries.

Historically, the word ‘timbre’ has its origins in Greek word “tympanon”, “timbanon” or “tumpanon” meaning ‘drum’. English language borrows the word from medieval French and there is not any record of the word used in any other meaning than some type of drum before nineteenth century. Dictionary sources indicate varying dates in the same century, Merriam-Webster’s dictionary points to 1845 for the first use for following definition: [T]he quality given to a sound by its overtones: such as (a) the resonance by which the ear recognizes and identifies a voiced speech sound, (b) the quality of tone distinctive of a particular singing voice or musical instrument. (Timbre, n.d.).

Despite its late use as a defined musical dimension, use of timbral aspect in composition was a developing idea since a few centuries, especially accompanying the developments in instrument-building. It had been customary to mention about Debussy and his harmonic language when talking about timbre, however, there are some examples of intentional timbral approaches that existed long before twentieth century.

European Art Music, like any other musical culture, had preferences about timbre perhaps since its oldest roots. Considering the divide between the musics of the people and The Church, ‘preference’ here indicates that they are either decided on a theological or philosophical basis. Roughly, last 1500 years can be seen as the larger era where we can follow those preferences and their effects, since it is a time period where Western Art Music is documented. The dominance of vocal music was gradually weakened as it developed into more complex understanding of harmony and greater structures during the Medieval and Renaissance eras. By Late Renaissance, vocal music had many genres, styles, and forms; and musical instruments and instrumental music was rapidly developing.

Baroque can be seen as the period where instrumental (thus, timbral) preferences really diversified and instrumental music began to become dominant in art music. One of the first notated and explained timbral approach can be seen in Jean-Féry Rebel’s ballet, *Les Élémens* (1737). Rebel does not use the word ‘timbre’ in his ‘warning’ (*avertissement*);[4] however, he openly describes his concerns on sounds and symbols.

[4] Related parts of the text in original spellings is as follows: «[…] *L'introduction a cette simphonie êtoit naturelle ; C'es toit le cahos même, cette confusion qui régnait entre Les Elemens avant l'instant ou,*

First movement of the composition, named "Le Cahos",[5] presents a depiction of chaos in nature before the "invariable laws" *(loix invariables)*. Firstly, he explains the relations between musical structures and gestures and the elements;[6] and he adds the following sentence to the cluster chord (Figure 1.1) of the first moment: *[...] J'ay hazarde de faire entendre d'abord tous les sons mêlés ensemble, ou plustout toutes les notes de l'octave réunies dans un seul son. [...]* (First, I took the risk to make all the notes sound together, or rather, all the notes of the octave united in a single sound.) (Rebel, 1737).

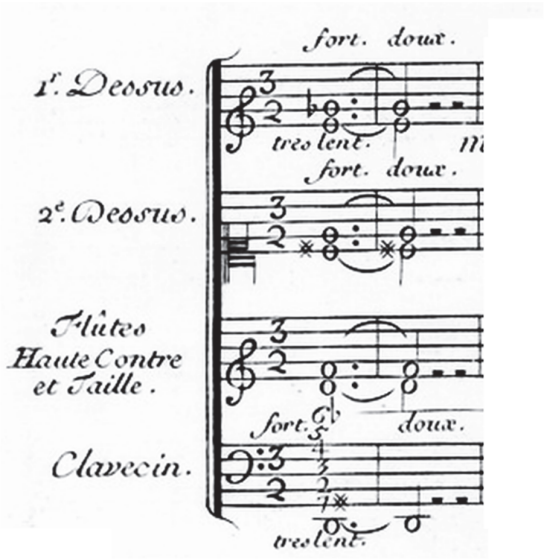

Figure 1.1 : Opening chord of "Le Cahos".

The initial chord including 'all the notes' from D minor scale then develops into "*l'accord parfait*" (perfect chord, in this case a D minor chord) with a rhythmic accelerando, as it is something "natural" to happen. Nevertheless, even if the first timbre-chord's position is taken as a negation of the 'order', this is still a proof of

assujettis a des loix invariables, ils ont pris la place qui leur est prescrite dans L'ordre de la nature. [...] Pour designer, dans cette confusion ; chaque Element en particulier je me suis asservi aux conventions les plus reçües. La Basse exprime La Terre Par des notes liées ensemble et qui se jouent par secousses ; Les Flutes par des traits de chant qui montent et qui descendent imitent le cours et le murmure de L'eau ; L'air est peint par des tenües, suivies de cadences que forment les petites flutes, Enfin les violons par des trais vivo et brillans représentent l'activité du feu.
[...] J'ay hazarde de faire entendre d'abord tous les sons mêlés ensemble, ou plustout toutes les notes de l'octave réunies dans un seul son. Ces notes se développent ensuite en montant a l'unisson dans la progression qui leur est naturelle, et, après une dissonance, on entend l'accord parfait. […]»

[5] Today, mostly spelled as 'chaos'.

[6] As presented: The Earth, The Water, The Air, The Fire.

timbral thinking. The expression of 'a single sound' is not far from any timbral approaches of twentieth century.

Even though there can be found some cases of timbral approaches/implications similar to Rebel's[7], they are not very common until twentieth century. Debussy, Ravel, Scriabin and in general, impressionist composers with their approach to chordal timbres contributed to the growing awareness. Later composers like Stravinsky, Varese, and especially Schoenberg with his concept of *klangfarbenmelodie*[8] (Schoenberg, 1978, p. 421) had timbrality in their music and discourse in following decades.

In the chapter titled "Non-Harmonic Tones" from his Theory of Harmony *(Harmonielehre,* first published in 1911*),* Schoenberg discusses the "harsh" and "ugly" chords that are avoided to integrate in the "system"; and he, in a way, defends these sounds saying "I maintain that these are chords: not of the system, but of music." (Schoenberg, 1978, p. 322) While he builds up a logic for timbral thinking in his own music, he uses the word *Geräuschmusik* (noise music) in an ambiguous way and says that "[…] there is no need to consider these, nor the many other things that appear in what is today called *Geräuschmusik*, 'noise music'[…] (p. 322). These comparisons exist and evolve throughout the Theory of Harmony. On the one hand he defends the validity and equality of these 'noises', and on the other, he indicates that all of these mean the system is collapsing.

As Brown (2014) points out, Schoenberg was referring to "harmonic others" (p. 111), the non-Western Art Music of the time; not only as the 'noise' as we understand today. However, it could well be included the new-born genre[9] of 'noise' if it were a decade later. Luigi Russolo's manifesto "The Art of Noises" *(L'arte dei rumori)* was published in 1913; and as Schoenberg's ideas represented 'the emancipation of dissonance', Russolo's represented 'the liberation of noises'. Russolo categorized noise types and devised notational systems and new instruments *(intonarumori)*[10] for

---

[7] Like Wagner's Prelude to Das Rheingold, which has a Eb major chord building up for almost five minutes.
[8] *Klangfarbenmelodie* (Tone-color-melody) will be revisited in later chapters.
[9] The term 'genre' may also be used as an inner categorization of Western Art Music based on their instrumental and conceptual characteristics (e.g., concerto, chamber music, vocal music, symphony, opera). In this thesis, it is used as an umbrella term for both these categories and all non-art music genres/styles (e.g., blues, pop, rock, folk, etc.) together.
[10] A set of noise producing instruments invented by Luigi Russolo.

the new genre. Following decades witnessed the development of noise music and its spread into many styles in different genres including many popular ones, industrial music, rock, post-digital music and many more, until today.

Noise instruments of Russolo disappeared but their timbral effects echoed in the following decades. Electronic music had its boost due the availability of many wider range electronic equipment and computers, especially between post-World War II era and 1980's. This period can be defined with many texture related musics. Perhaps the most important one in the subject of genre was Spectralism; and it arrived together with the simplistic and focused nature of minimalism, and the advancement of computer-based sound analysis.

The term 'spectral music' was first used by Hugues Dufourt in his article *"Musique spectrale"* in 1979 (Dufourt, 1991), to define a rising compositional way of thinking in contemporary music. However, its roots are very often linked to the compositional tradition which started with Claude Debussy and his understanding of sound as a perceived object and focusing on the acoustic qualities of every instant. During the same period, George Enescu composed possibly one of the first proto-spectral pieces for piano (Szász, 2011; see 2.3.1). This tradition continued with Edgard Varèse, Giacinto Scelsi; and especially Olivier Messiasen, who either directly taught or became an inspiration for the first generation of spectralists. However, apart from other Europeans, perhaps Per Nørgård's[11] 1968 work "*Voyage into the Golden Screen"* was "the first properly instrumental piece of spectral composition." (Anderson, 2000, p. 14) Actually, there was no real school of spectral music or a group of spectral composers (Anderson, 2000), but only composers who were interested in spectral features of sound, in order to find new timbres; or who use these basic scientific features as a source of inspiration for new compositional techniques. These composers can be found worldwide, but the main impact to the field was made mostly by French and Romanian composers with their varying uses of the same phenomena.

Spectral compositions were mostly built using harmonicity-inharmonicity contrast (See 4.2.5), interpolations, modulations. Most importantly, with the availability of analysis and synthesis techniques, the notion of timbre became the main inspiration, the working material, and the result in spectralism. Spectral thinking was used to create

[11] Per Nørgård (born 1932), composer from Denmark.

new pitch systems based on the overtone series and to produce a higher level of timbre. And at the same time, the sound-world realized through these new possibilities has deeply affected the aesthetics of contemporary music. From Debussy's timbral view of chordal melodies, it had come to a point that timbre was in relation with almost every aspect of music.

Today, timbral approaches and techniques can be found in many different styles of Western Art Music compositions. The ‘discovery’ of timbre as a compositional tool has succeeded to new level, where its scientific and philosophical features are researched and applied in greater content. Therefore, following chapters will cover the subject of its cognition and its relation with main structural elements of music.

2. TIMBRE, ITS COGNITION AND CATEGORIZATION

Contemporary music, regardless of genre, has become more complex in timbre (and in other aspects); and it requires to be analyzed in more dimensions, in connection with many factors. In many practices of composition, there could be inconsistencies between the goals at the compositional stage and the result at the final stage where the listener's perception of the work is the actual result. Compositional material can get steadily more complex while the composers think they have created something completely new for the ears. Yet, it is always limited with the listener's perception.[12] Hence, many composers and researchers of mid-twentieth century were concerned with the reception of the extremely complex modern music. Some of them reacted with "call for a return to tonality", and some suggested a scientific approach to music, through the study of human perception (Born, 2012, p. 420).

The nature of timbre was researched through many studies since a few decades. One of the important long-time researchers on timbre, Stephen McAdams, published studies in collaboration with many colleagues, mainly working on multi-dimensional space (MDS) of logarithmic attack time, spectral centroid[13] and spectral flux[14]. Through these experiments, they mapped an example of timbre space of many sources, varying from sine wave to real instruments, with the addition of synthesized hybrid instruments (McAdams et al, 1995; McAdams, 1999). This research underlined the necessity of considering timbre as a multidimensional property of the sound.

Using these basics for their experiments, many other scientists came up with new findings about timbre and its cognition. Early studies mostly focused on the lateralization of timbre in brain because it could be an indication of how timbre is

[12] Of course, there is nothing essentially wrong with creating something that is already shaped by any tradition. Putting something that is not completely new or using allusions is not an artistic crime. However, it is important to be sure about whether it is new or not in the listeners' perception.
[13] Spectral centroid: The center of frequential density in a spectrum.
[14] Spectral Flux: The rate of change in the sound spectrum.

processed.[15] Results were indicative of a higher usage of right hemisphere during listening of musical sounds. However, more consistently with the proposed multidimensional approach, some later functional magnetic resonance imaging[16] (fMRI) studies proved that both left and right temporal lobes were active during timbre perception (Meyer et al, 2006, p. 1511), with an asymmetry concentrated on the right side. Meyer et al (2006) found that there was significant increase of amplitude in N1 and P2[17] when instrument sound was given as stimuli, compared to sine waves (Meyer et al, 2006, p. 1517). During N1 time window, they observed density changes in primary and secondary auditory cortex and anterior insula[18] for both; and during P2, low-resolution electromagnetic tomography (LORETA) showed increased density distribution in auditory cortex bilaterally and in the frontal gyrus; which indicates higher cognitive activity, such as emotional and auditory imagery functions (Meyer et al, 2006, p. 1510). Another recent research that used instrumental tones and sine wave showed that the involuntary processing of timbre variations were more related to discrepancy between the standard and deviant tone rather than the spectral complexity (Lai et al, 2011); which supports not only previous findings but also indicates the significance of relativity factor (See 2.3.3) in timbral perception.

2.1 Timbral Continuum

The term 'timbral continuum' is used by many scholars before, to indicate similar but separate things. Here it simply describes possible dimensions of timbre that represent graduality in varying degrees. This notion can be explained in two topics: physical (acoustic) and perceptual; and the perceptual side of the continuum is the main element of the timbral coherence factor.

---

[15] It is often accepted that left brain functions as the analytical side (logic, mathematics, language, reasoning etc.) whereas right side as the imaginative side (creativity, art appreciation, music etc.) However, there are many studies that show these are not completely separated functions, rather relatively concentrated on either side.

[16] Functional magnetic resonance imaging or functional MRI (fMRI) measures brain activity by detecting changes associated with blood flow.

[17] N1 and P2 (Negative 100 and Positive 200) are event-related potential components which are active during sensory and cognitive stages of auditory process. Here 100 and 200 stands for miliseconds. N1 is the negative peak after 60 ms following the stimulus and P2 is the positive peak between 100 and 200 ms.

[18] Recent studies associate the anterior insula with cognitive choices and intentions, music, time perception, awareness of sensations and movements, of visual and auditory percepts besides many other functions.

A significant explanation of physical continuum of timbre can be Denis Smalley's suggested spectral typology (Emmerson, 1986, p. 65). This proposal is structured as a continuous spectrum between “note proper” and “noise”.

Note proper is mainly the subject of the pitch-based music traditions that use the intervallic combinations and conventional performance techniques. Therefore, the output of note proper music is more about the conventional meaning of harmony, than spectral quality. Because theoretically, note proper can be seen as an indicator of pitch, more than a real sound.[19] Two other categories under ‘note’ are instruments/sound objects with harmonic spectra and with inharmonic spectra and in fact, harmonic and inharmonic spectra can be present in note proper. The second parallel dimension is the note to noise continuum and the noise spectrum is where no internal pitch structure is perceived.

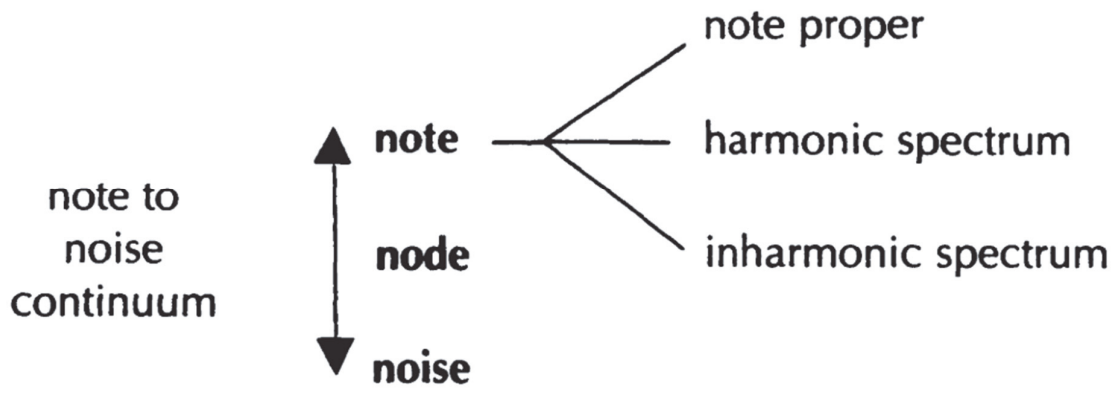

Figure 2.1 : Spectral typology of Smalley (Emmerson, 1986).

Tonal music usually focuses on only one spectrum at one moment;[20] and in a simplified form realized through note proper. Additionally, the perception of the existing simplified spectrum is distracted with conventional gestures. Thus, timbre exists and is perceived only through separate instruments;[21] the complete spectra[22] of the elements is ignored. Therefore, perceptual side of timbral continuum cannot exist

[19] But at the same time, a pitch written for an instrument indicates a specific timbre (See 2.4.1, ‘Macrotimbre’).
[20] Spectral features of certain chords/combinations can be discussed, but the goal of tonality is always focused on chords from a selection of certain timbral qualities.
[21] This is especially relevant in the case of contrapuntal music. See 2.2, ‘spectral fusion’.
[22] See 2.4.1, ‘overall timbre’,

on the awareness level. It is possible to see the harmonic progression as a layer of timbral change; however, it could be a limited one.

Perceptual side of timbral continuum is dependent on several other factors. Specifically, perception of timbre changes in relation to the effect of density of the musical information. Thus, temporality aspect creates a continuum between timbre and others. One of the concepts proposed for this transformation is called "apperceptive modulation" by Teodorescu-Ciocanea (2003, p. 97). The apperceptive[23] modulations concerning timbre can be listed as:

> 1. 'Harmony/Timbre Apperceptive Modulation' is the change in perception as a result of increasing number of pitches in a chord. When the chord exceeds the limit of pitches that can be perceived separately, it becomes a complex chord, which is perceived as a timbre. This modulation can be seen as one of the essentials of spectral music, since most of the chords are presented in large quantity of pitches. This type of modulation is partly described before by Schoenberg (1978) in *Harmonielehre* chapter named "Aesthetic Evaluation of Chords with Six or More Tones" (p. 411) where he also mentions the notion of *klangfarbenmelodie*.
>
> 2. 'Rhythm/Timbre Apperceptive Modulation' is the transformation of rhythmic elements into timbre through different levels of repetition.[24] Thus, the gap between events gets smaller than we can perceive. Through this modulation, pitched elements can gain a new level of timbre and non-pitched ones can have pitch, therefore timbre. Although this modulation is not an essential one, it still can be observed in many works as a supporting gesture.
>
> 3. 'Timbre/Form Apperceptive Modulation' occurs when density of the timbral information (both horizontal and vertical) increases and exceeds the limit of perceivable sonic events. Listener's focus shifts towards the organic sound mass which, with the help of temporality, turns into an element of form. This

---

[23] Apperception (in psychology) is "the process by which new experience is assimilated to and transformed by the residuum of past experience of an individual to form a new whole." (Runes, 1972, p. 15)

[24] Karlheinz Stockhausen's 'Kontakte' (1958-1960) is a well-known work that uses this phenomenon. Stockhausen himself explains how he built the work through these concepts in his 1962 article titled *"Die Einheit der musikalischen Zeit"* (The unity of musical time). English translation by Elaine Barkin is titled "The Concept of Unity in Electronic Music".

type of modulation can be most practically evident in mimetic spectralism, which uses sound analysis and modelling techniques; however, this thesis takes this in relation with the general understanding of the musical form and work based timbral identity.

These modulation types are essential to comprehend the approach that is going to be presented in later sections. Interpretations and adaptation of them are needed from micro to macro level.

2.2 Vertical and Horizontal Organization in Composition

Timbre has an essential role in the perceptual organization of musical structures. This organization mechanism is made up mainly of auditory stream integration and auditory stream segregation (See 2.3.4). These processes use the spectral and temporal data to connect/separate sound events and create continuous streams which form more explicit musical structures.

Firstly, the vertical organization starts with the analysis of the consistent partials forming one timbre, which will be perceived as one sound source. This consistency keeps timbres of the instruments separate even if they have very close harmonics in upper registers. On the other hand, the same effect makes some instrument combinations blend better than others, depending on their spectral features.

Secondly, starting moment (or onset) and the behavior of partials over time are unifying factors too. So, every entrance of single pitch in a timbral chord sums up the previous ones into one perceptual object (Also related to Gestalt, principle of proximity, See 2.3.4). This entrance could be any strong gesture. One distinctive example for this type of grouping can be seen in Gérard Grisey's work, *"Vortex Temporum"* (1995) (See Figure 2.2), where rapid gestures of some instruments are triggered by sharp attacks in order to combine them spectrally.

Lastly, the location of the sound source helps organizing verticality. It is quite common to have spatial designs with this approach, and we can find some examples of this concept in both acoustic and electroacoustic works. However, it still could be debatable whether localization of a timbre created by an ensemble or orchestra would lessen the intended effect of timbral composition of sound.

On the practical side, due to perceptual factors above, vertical organization is double filtered during performances. Because the performer, most of the time, is the first to perceive the spectral structure of the moment and thus, will able to fuse it better. Since the same process takes place in listener's perception, spectral accuracy becomes less questionable. Nevertheless, some factors like practical difficulty of musical gestures and tempo still can affect this in the opposite way. In addition, in any case of re-listening of a certain timbral structure (by a performer or non-performer) provides a layer of objectivity, providing a better auditory analysis in perception.

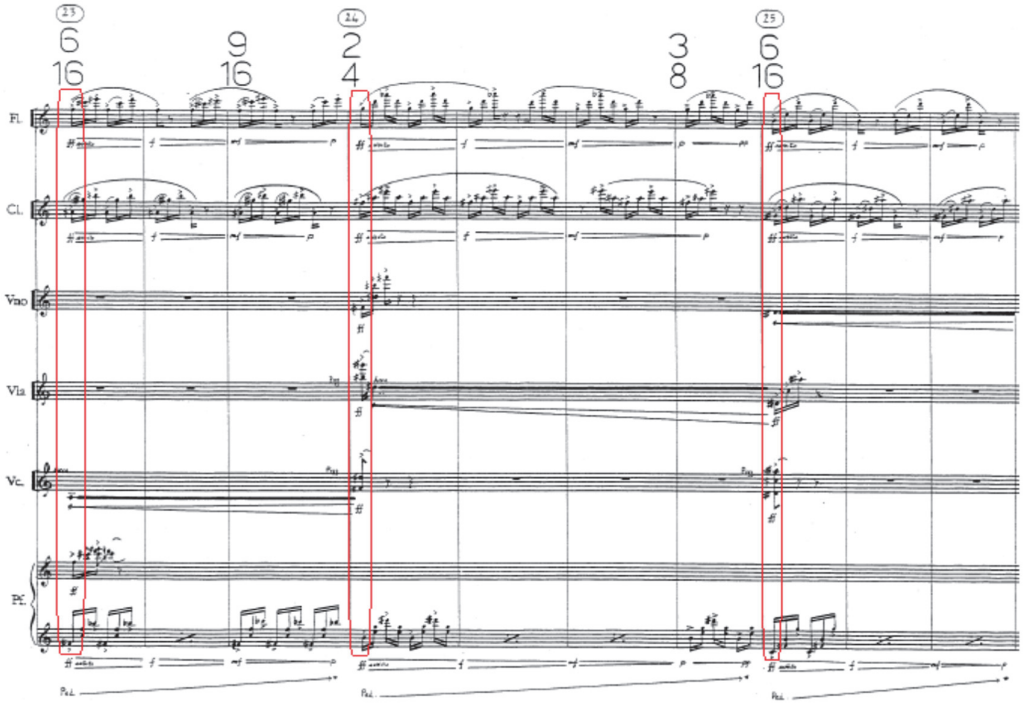

Figure 2.2 : Excerpt adapted from "Vortex Temporum" by Gérard Grisey.

Horizontal organization is the grouping of musical structures which takes place in the temporal dimension. Spectral continuity of timbre helps keep events perceived as coming from the same source. For example, through horizontal organization ability, we can choose and follow the melodic line of a certain instrument in an orchestral setting. Besides, if there is a continuous timbral change in the sound source, it is possible to perceive it as one stream horizontally. However, perception of this kind of timbral change has to be directed (See 2.3.3).

An alternative example of organization and spectral fusion can be found in Murail's *"Désintégrations"*. Seventh section of the work is based on an inharmonic spectrum,

which is timbrally more ‘disintegrated’ (See Figure 2.3). He uses rapid parallel movement to fuse the sound into one timbre. The inharmonic nature of the chords is transformed into material that can be analyzed by pitch-melody related approach. Thus, he uses vertical and horizontal perception against each other; yet, from a compositional perspective, they work together to form a certain timbral region.

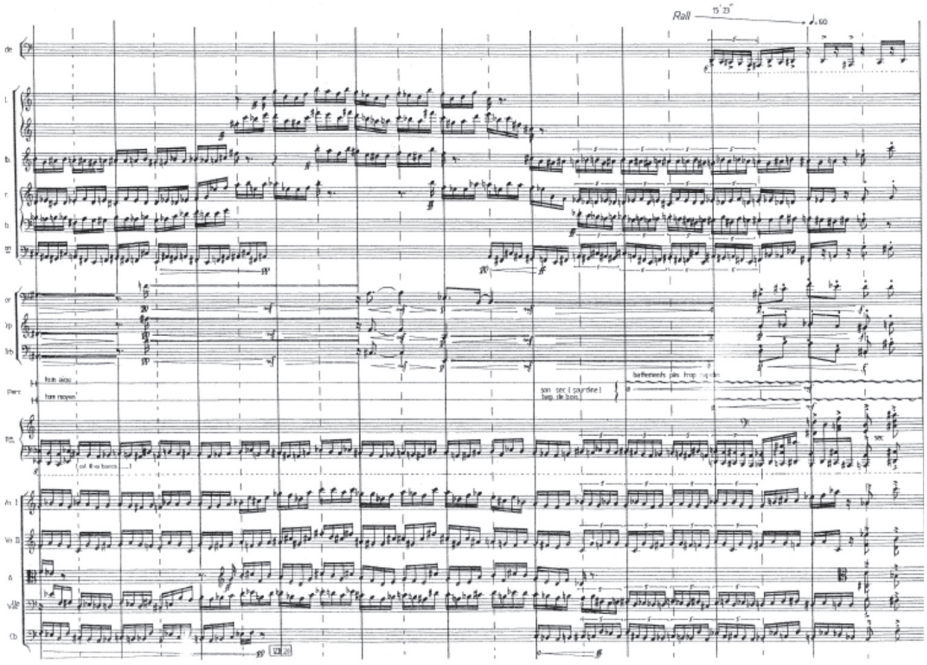

Figure 2.3 : Excerpt from the seventh section of *“Désintégrations”* by Tristan Murail.

In addition, the strong gestures that can act as timbrally unifying factors (as stated above), would also be a distracting factor that breaks the horizontal continuity. This way, horizontal continuity of one line hides the continuity of another horizontal evolution of a timbre. Wright and Bregman (1987) studied this subject and stated that “vertical and horizontal forces compete with one another” (p. 73). This competition means spectral fusion works against horizontal forces of streaming. So, avoiding one dimension, could enhance the perception of the other. Similarly, Huron (1991; 2001) showed that in Bach’s music spectral fusion[25] was avoided. He stated that “Bach preferred intervals in inverse proportion to the degree to which they promote sensory

[25] Same concept is used as “tonal fusion” in the mentioned article.

dissonance and in inverse proportion to the degree to which they promote tonal fusion" (Huron, 2001, 21).

2.3 Horizontality and Melody

A further investigation of timbre requires a topic related to one of the main structures in music: melody. This investigation is crucial for this thesis, considering the fact that melodic structures are the most common elements that shape the gestural bodies (which will be examined in Chapter 3).

To begin with a formal definition, melody can be defined as a group of pitches that are sequenced in time and have diverse properties depending upon the musical culture that they originate from. It is one of the oldest and most common aspects in music traditions around the world. It was the focus of Western Art Music tradition for a long time and it still occupies an important role in many contemporary musical genres. Basic aspects and principles to comprehend the notion of melody can be listed as follows:

- Pitch is the main unit for defining melody. Here pitch indicates the synthetic notion of an interval relations and perceived fundamental of any musical sound.
- At least three pitches are needed to define a melody.[26] One pitch is a simple reference, two notes indicate a direction, three notes can form a closed Gestalt.
- Melody is perceived retrospectively; it is closed as a unit/group depending on temporal positioning (inner [rhythm] and outer) and tradition-based interval relations. Thus, melody is a gesture, defined in vertical and horizontal aspects.
- Timbre is a second layer that would be necessary for the identity of a certain melody (See 2.3.5).

Regardless of genre, melodic connotations can occur almost anywhere. As a combination of gestural and timbral properties, this mid-ground temporal element has always a potential to contribute to form. Therefore, it is necessary to take a look at some of the significant studies on this topic and compare their findings. Firstly, interaction of individual sub-topics of melody will be covered.

[26] Two pitches would serve as a 'figure' rather than a melody. See 3.2.2

2.3.1 Interaction of timbre with pitch and dynamics

There have been many studies on the relation of pitch perception and timbre since a few decades. These experiments enabled us to understand that the spectral composition of the sound was a source for timbre, pitch, and the interaction between these two. These studies had to improve their tools until today.

Pitch is the most apparent element of a melody alongside rhythm. However, timbre is not less relevant. An important and very clear statement about this relation by Schoenberg in the last pages of his third edition of *Harmonielehre* from 1922 is as follows:

> The distinction between tone color and pitch, as it is usually expressed, I cannot accept without reservations. I think the tone becomes perceptible by virtue of tone color, of which one dimension is pitch. Tone color is, thus, the main topic, pitch a subdivision. Pitch is nothing else but tone color measured in one direction. (Schoenberg, 1978, p. 421)

Furthermore, in following lines in the same paragraph, he uses the term 'coherence' (*Zusammenhang*) when identifying melodic and harmonic progressions; and says that it would be possible to apply the same thing for 'tone colors' (*Klangfarben)*, following the logic to the notion of *klangfarbenmelodie* (Schoenberg, 1978, p. 421).

Timbres of musical instruments change along with their pitch (See 2.4.1, 'macrotimbre'). Most instruments are easily identifiable when compared within an octave (McAdams, 2013, p. 45). Some studies which use greater pitch intervals between two instruments, have found out that listeners were either ignoring the pitch difference; or pitch differences were physically creating timbre differences due to the change of spectral centroid (McAdams, 2013, p. 45).

From the musical perspective, this finding points out two important things. Firstly, involuntarily ignoring the pitch/timbre difference provides an advantage in building more homogeneous timbre chords (i.e., elements can be perceived as the members of the same continuum). Secondly, it requires a greater research to combine instruments in smaller groups, when any spectral continuity is sought between those.

Dynamics also create timbral change in instruments. The spectral centroid may change depending on the energy applied to the instruments. Parallel to the dynamics from pianissimo to fortissimo, the centroid of the spectrum moves to a higher band, making the timbre 'brighter' as in the common use of the word (McAdams, 2013, p. 46). Thus,

a chord played piano (***p***) or forte (***f***) (with an ensemble or orchestra) will not have the same timbre even if all instruments are in the same dynamic balance. This means the phrase 'same dynamic balance' becomes essentially impossible. Therefore, a slightly higher balance in pitch is required with softer chord and the opposite with louder ones.

An early and significant compositional application of pitch, dynamics and timbre relation can be found in George Enescu's *Carillon Nocturne* (1916) from his Piano Suite No. 3. Enescu imitates the timbre of a carillon[27] by organizing the vertical elements in different dynamics and articulations and synchronicity.

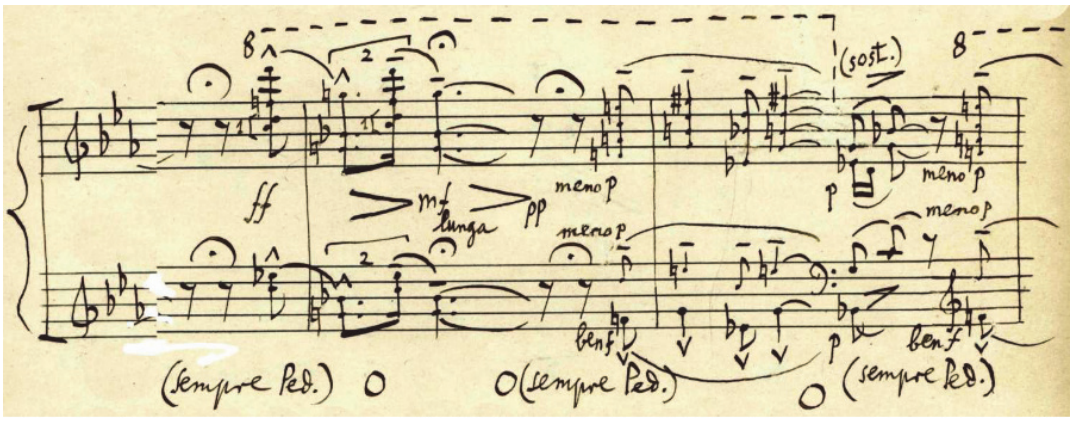

Figure 2.4 : Excerpt from the original manuscript of "Carillon Nocturne" by Enescu (Szász, 2011).

In the original score (Figure 2.4), lowermost notes are 'large notes' and others are 'small notes'. Large notes are played with slightly stronger dynamics and small notes are with softer. In addition, Enescu uses three main intervals to create a certain timbre through the chords: "the melodic major third, "cloned" major sixths clusters, and virtual augmented octave false relations" (Szász, 2011, p. 8). These intervals are used in parallels, in a Debussyan[28] way. By doing so, he creates a certain quality of timbre through imitation. However, despite the general effect of sustain pedal, the lower 'strike notes' also differ from smaller 'tenuto' notes in sustain portion. Totality of dynamic balance and note-head differences can be seen clearly in a 1958 edition of the same score in Figure 2.5.

[27] An instrument that consists of bells that are produced to be able to be played melodically. Usually installed on a church roof, played by an organ-like remote mechanical system.
[28] In a sense that is 'similar to Debussy' and not 'post-Debussy'.

This type of interaction will be contextualized in Chapter 4, where the registral and spectral properties are examined in an analytical approach.

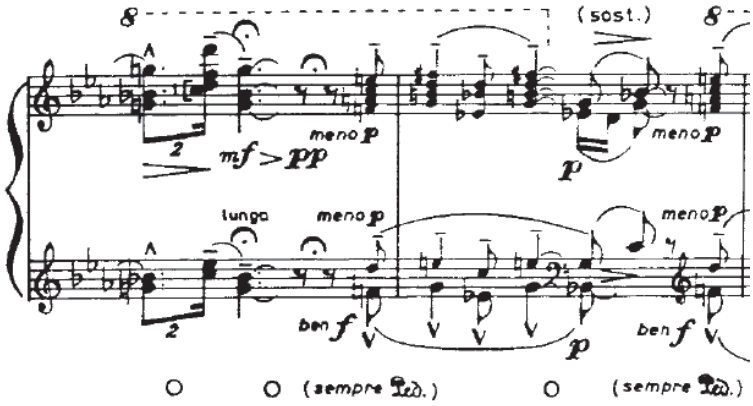

Figure 2.5 : Excerpt from the 1958 edition of "Carillon Nocturne" by Enescu.

2.3.2 Melodic intervals and intonation

Russo and Thompson (2005) studied this interaction using digitally synthesized sounds with different balances of partials in order to get 'dull' and 'bright'[29] timbres. They chose two intervals, the tritone and the perfect fifth; and they played these sounds melodically (with 1 second durations each) to get three different types of timbral change: congruent (ascending interval getting brighter, descending interval getting duller), incongruent (ascending interval getting duller, descending interval getting brighter), or without any timbral change.

Their experiments had some significant results. Firstly, they showed that a tritone (six-semitones) interval with a congruent timbre change could be perceived larger than a perfect fifth (seven-semitones) interval with an incongruent timbre change (Russo and Thompson, 2005, p. 559). Russo and Thompson's conclusion basically means that the effect of change in spectral centroid could be so strong that it could affect our interval perception, without any change in perception of the pitches separately. However, the participants in the first experiment had two years or less of music education, therefore, they wanted to repeat the experiment with musically trained (10 years or more) participants. Results of the second experiment showed that "interval illusion" existed

[29] Brightness can be defined as the attribute of sound which is a result of the placement of spectral centroid (or frequential gravity center). It is one of the most perceivable qualities of timbre; and most of the timbre related studies select brightness as a variable for their experiments.

among trained musicians too, but with a difference that it was more apparent in descending intervals rather than ascending. These experiments are concluded with a final one which re-checked the possibility whether there was any change in the perception of tones; and it confirmed that these timbral manipulations did not affect the pitch, but the perceived interval size only. Considering the relative importance of these intervals musically, this result alone suggests us to reconsider our understanding of interval analysis in different instrumental combinations. Nonetheless, an experiment with synthesized sounds may not be convincing enough.

In a more recent study, Vurma et al (2011) used real instrument sounds and avoided any possible criticisms based on this aspect of similar studies. This study concentrated on two questions. First was whether the timbre difference between instrument sounds (and voice) cause a perceptual pitch-shift with a significance that would affect intonation quality; and second, if such a shift exists, whether it is dependent upon musical education of the listener, as investigated in Russo and Thompson study. For these tests, they used recorded sounds of a (bowed) viola, a trumpet and a singing voice which are based on the same fundamental frequency (F0)[30]. Sounds were played with two seconds of silence in between. Participants, an equal number of musicians and nonmusicians, were asked to rate pairs of successively played sounds as 'in tune', 'sharp' or 'flat'.

The experiment confirmed the expectations from previous ones and showed that there would be approximately 15 to 20 cents of difference in F0 perception between different instruments; and even 8-10 cents of difference between same instrumental sounds. Results were also consistent with the conclusions from Russo and Thompson's study where they found out that F0 of the sounds with brighter timbres were perceived as higher and the ones with duller timbres were perceived as lower. These results are also consistent with Terhardt's approach on virtual pitch, or missing fundamental (Terhardt, 1974). According to Vurma et al, Terhardt's algorithm to calculate the virtual pitch indicated that perceived F0 of the trumpet should be 16 cents higher than the viola, which was confirmed by this experiment.

After evaluating these results, it is still a question if these pitch differences could affect the intonation, thus the quality of the performance or quality of listening. Firstly, we

[30] F0: Fundamental frequency, referring to the first member of a harmonic series.

know that such differences (no more than 20 cents) do not create perception of any separate pitch for most people since they fall in critical band.[31] But this doesn't mean that they do not make any difference at all. This amount of difference is audible and creates a difference in timbral quality. In another study, Sundberg, Prame, and Iwarsson (1996) compared critical assessments of certain recordings containing vocal performances and their actual F0 differences found out that the acceptable range (by the experts) was no more than ±7 cents (As cited in in Deutsch, 2013, p. 97).

In conclusion, melodic intervals and timbre relation can affect the timbral quality in certain conditions, and these conditions can occur regardless of the level of musical training. More detailed studies on other instrument combinations are needed to explore the map of possible results of this interaction.

2.3.3 Sequences and relativity of timbre

The next step to get to the conventional meaning of melody is the concept of sequence (as a set of elements in a particular order). Through examination of some studies on sequences, we can find out possible hierarchical relations between timbre and melody; and perhaps, question the relativity or absoluteness of timbre. In a 1992 study, Krumhansl and Iverson questioned if there is a proof of interaction of pitch and timbre. Depending on the results of the first experiment, they concluded that the participants were not able to attend to the pitch of a tone without being influenced by its timbre and vice versa; but in these limits, they were still able to recognize same pitches and same instruments when compared.

Their second experiment focused on pitch and interaction in sequences. Using two variables (pitch and timbre), they prepared a three-level setup where they respectively changed (i) a certain pitch in the sequence, (ii) the complete pitch order and (iii) timbre order with one pitch change. The changing pitch (the target) was always at the same place in the sequence. By using this method, they came up with a quite important conclusion: "changing the pitches of tones surrounding the target tone did not interfere with memory for the timbre of the target, and changing the timbres of tones surrounding the target tone did not interfere with memory for the pitch of the target

[31] Critical band here is the limit of frequential difference at which our ear can separate two very close tones. Frequencies belonging to the same critical band fall activate the same region in the basilar membrane, where all frequencies are processed in the ear.

tone" (Krumhansl and Iverson, 1992, p. 747). This finding indicated a level of independence for these two aspects. However, the same experiment also indicated that changing pitches of the tones affected the perceived pitch of the target tone but the same finding was not valid for timbre, which means that the perception of a certain pitch always depended on the relation of other pitches, but timbre was independent through whole process.[32] This result points out that timbre may be more absolute than we think.

The third experiment by Krumhansl and Iverson concentrated more on timbre and its perceptual relativity. A similar group of constant and varied timbre sequences was used with pitch sequences. Results supported the previous findings: timbre changes in a pitch-varied sequence had almost no effect on pitch perception. Timbre perception was always worse at any condition with a pitch change. Their article also included the comment that a better perception of timbre patterns required high control of pitch. To conclude, timbre perception got better only when there was less attention on pitch and these experiments showed that relative perception of timbre was unstable. In contrast, relative pitch perception was a "robust phenomenon" (Krumhansl and Iverson, 1992, p. 749).

This relative weakness of timbre against pitch can be investigated through some other connections. A 2014 study by Cousineau, Carcagno, Demany and Pressnitzer does it from a different perspective. Their article starts with the aforementioned quotation from Schoenberg's *Harmonielehre*, in which he, consistently with his *klangfarbenmelodie* idea, describes the pitch as "nothing else but tone color measured in one direction."

In their experiments, they chose brightness as the timbral variable. In addition, since the change in timbre is about the change of amplitude in certain partials of sound, they included loudness sequences along with pitch and brightness. So, the main question was whether timbral sequences are processed as/similar to pitch sequences, or loudness.

The results of the experiments revealed that pitch sequences were processed more efficiently than loudness sequences. Increasing complexity of sequences was less

[32] For related Gestalt principle, see 2.3.4, 'good continuation'.

affective in pitch sequences than in loudness sequences, but the surprising result was that brightness sequences were "indistinguishable" from the pitch sequences (Cousineau et al, 2014, p. 5). This result is important since (as stated before) the changes in brightness was completely dependent on amplitude change of harmonics (not elimination or addition of them) just like loudness change. On the other hand, pitch sequences were dependent on frequency change of a complete set of harmonics.

Even though these results are consistent with the findings related to studies on melodic intervals that show the influence of brightness on pitch perception, they still cannot present an explanation for the relativity problem. Nevertheless, Cousineau et al tried to look for other explanations, like the possibility of a common cognitive mechanism related to the storage, use or retrieval of pitch and brightness. However, they also indicated that there is no evidence for this kind of hypotheses yet, and it remains "largely mysterious" (Cosineau et al, 2014, p. 14).

A possible temporary conclusion can be drawn with the inclusion of McAdams' studies (McAdams and Cunible, [1992]; McAdams et al [1995]) on timbral intervals, which this more recent article uses as a reference. It is mentioned that "timbral intervals" are perceivable but "without singling out the brightness dimension and with large inter-individual differences" (Cosineau et al, 2014, p. 5).

McAdams studied the topic since his collaboration with Cunible in 1992; and in this study he tested a vector model where he used Krumhansl's tri-dimensional space from 1989 (Actually, another very similar timbral analogy test was already presented a decade ago, by Wessel in 1979). [33] The main idea was to create a multi-dimensional model to explore perceptual space of timbres where they could (possibly) be studied in a similar way as it could be done to pitch intervals. These works used this tri-dimensional space with synthesized instruments and their hybrid forms. In his conclusions on timbre intervals in 2013, he pointed out the risk that timbral intervals contain: real life instruments have many specificities which will not allow any kind of application. 'Specificity' here means the unique attributes to any timbre; e.g., despite the harpsichord is commonly known for its sharp attack and bright quality in strings, the sound that is created by the returning hopper when the keys are depressed is a

[33] See 2.3.4 for more of Wessel's study (1979).

specificity.[34] It is one of the essential terms to understand the 'absolute aspects' of timbral qualities.

Krumhansl and Iverson's discussion on the absoluteness of timbre is very much related to the McAdams' conclusion about timbral intervals. As they stated that "whether timbre is perceived in absolute or relative terms is a bias that stems from considering timbre from a musical point of view" (Krumhansl and Iverson, 1992, p. 750).

Depending on the comparison of these related studies, this relation can be summed up in two topics. Firstly, scientific data created by mentioned scholars show that regardless of its solid existence and application in music, timbre is poorly relative in perception. The adjective 'absolute', that Krumhansl and Iverson use, is more likely to be applicable when timbre is the subject. Apparently, pitch is too simple, and its range is too narrow in comparison with timbre. That is why it has an evolutionary advantage in musical structure. Moreover, this change happens around one or two-year-old infants while they start to facilitate language learning, indicating it is a language related capability (Mithen, 2007, p. 276). In addition, there is also a debate that if in a certain stage of human evolution perception of pitch has been absolute and it is replaced by relative pitch later and if it is possible to train adults to gain absolute pitch (Deutsch, 2002, Wong et al, 2019; Carden & Cline, 2019).

To clarify, these studies show that pitch is perceived relatively; it –almost always– needs a strong reference to its neighbors to be perceived in a melodic context. However, if we consider the range of frequencies we mostly use for melodically used pitch registers (practically, it does not even cover the whole piano keyboard), timbre could be seen as a huge bulk of data for us to analyze, since it covers a greater register of frequencies and includes many other specificities that we are even not accustomed to think about.

Secondly, we should take cultural factors into account. Our conditioning in music is very strong towards certain aspects of it (in this case, melody), and for most of us it is quite difficult to hear the music from different perspectives. Aforementioned studies do not say much about this side of the subject. There still may be some factors in pitch and timbre cognition which are universal, dependent on our biological reality; but it

[34] Same example is given in McAdams et al, 1995.

would be wiser to consider the power of our exposure to (any kind of) culture, since some decades we have theories how genes left their primary function to the memes in human evolution.

Finally, it would be good to remember that the Schoenberg quote from his *Harmonielehre* contains a speculative detail. As Cramer (2002) states that early in the book, Schoenberg argues that "all the partial tones should be considered 'suitable for art'" and he says that "higher partial tones are perceived only as *Klangfarbe*, while the lowest partials are more 'familiar'", which are perceived as pitch. However, at the end of the book, he considers all tones as *Klangfarbe* (p. 10). This can be a result of the opposite case compared to the idea above: a case of self-enculturation; meaning, Schoenberg (as most professional musicians/composers) may have been a victim of his own ideas as he was more and more exposed to them; and this can be analyzed through the same perspective that apperceptive conditions occur (as explained before). There are studies focusing on effects of enculturation on timbre that suggest changes in dimensionality of perceptual timbre space (Vinoo and Toiviainen, 2012), however, not any on self-culturation among composers, yet.

2.3.4 Streaming, grouping, segregation

Perception of auditory streams is explained through "auditory scene analysis"[35] theory, which had its roots in Gestalt theories from the previous century (Cross, 2007). Gestalt theory claims that our perception is conditioned to find patterns in the data that we gather by our sensory systems, which makes the sum of the perceived more meaningful than the parts; and these patterns can be formed using many principles such as proximity, similarity and good continuation. For example, an instrument playing in a very high pitch or high amplitude can be segregated from others by proximity of frequencies or amplitude. In the case of timbre, we can easily imagine that an instrument with a brighter timbre can be identified among other instruments with darker timbres. An illustration of some basic Gestalt principles can be seen in Figure 2.6.

However, the scene analysis for pitch and loudness is easier than it is for timbre. The specificities of any timbre can create big differences in any auditory data, and it may

[35] Term coined by Albert S. Bregman in (Bregman, 1990).

be difficult to predict which of these would be the key to perception of an individual timbre. Yet, timbre functions as one of the most important factors to comprehend the Gestalt in music.

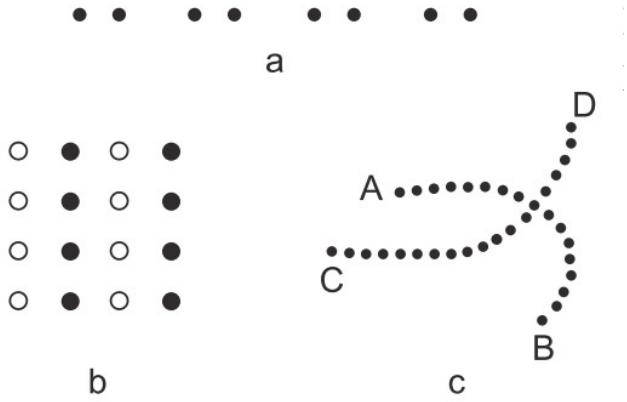

Figure 2.6 : Illustrations of the Gestalt principles of (a) proximity, (b) similarity, and (c) good continuation (Deutsch, 2013).

First of all, it is important to understand the amount of perceptual data that timbre loads in our perception. Deutsch (2013) presents the example of an experiment by Warren, Obusek, Farmer and Warren from 1969. Experiment uses four timbrally unrelated sounds (high tone, hiss noise, low tone, and a buzz-square wave) which were presented repeatedly to the participants. When the sounds had the duration of 200 ms, participants were unable to perceive the order of the presented sounds and when durations were raised to 500 ms, they were able to order them correctly. Referring to previous chapter, we can remember the significance of using instruments sounds rather than synthesized ones. Furthermore, aforementioned study by Meyer et al (2006) showed that there was significant increase of amplitude in N1 and P2 when an instrument sound was given as stimuli, compared to sine wave (which is 'theoretically' timbre-less)[36] (p. 1517). Today, we know that 200-450 ms could be the minimum amount of time to identify a timbre correctly and this is related to N1 and P2 which the components were defined at a later time than these experiments (Siedenburg et al 2019, p. 75). The results of these studies imply two possible conclusions: (i) These

[36] Naturally, it is not possible to get a sine-tone in the pure form. Once it is presented through a medium, it gains a spectrum through the physical properties of the medium and the instruments it is channeled through. An early example of this dilemma can be seen in (Toop, 1979, p. 387).

time limits can be seen as an indication of the amount of data that is stored within timbre; and (ii) if we are able to perceive/identify the timbral data (as a whole, not in an order), we still need more time for processing other aspects of the sound in order to perceive it in detail.

Depending on these aspects of timbre, streaming by timbre can be understood better. Wessel's 1979 study on timbre space included an example of how timbral properties can be used to form melodic lines. Participants were presented repeating patterns of ascending lines that consist of three pitches. The timbres of these pitches were alternating (See Figure 2.7) and the perception of groups was affected by the tempo they were presented in.

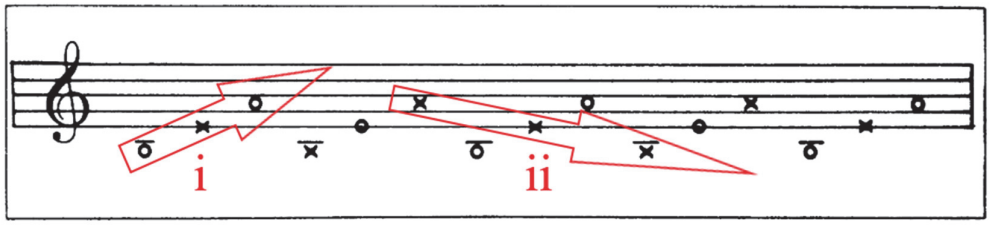

Figure 2.7 : "Ascending pitch patterns in 'three' with two alternating timbres ('O' and 'X')", adapted from (Wessel, 1979).

As a result, perception of lines changed with the amount of timbral difference between "O" and "X". When difference was small, the whole was interpreted as ascending lines (i), and when it was bigger, participants started to hear descending lines (ii) (Wessel, 1979). This is an example of grouping through temporal proximity in the first condition, and segregation through timbral similarity/difference in the second one. However, it is worth mentioning that another Gestalt based principle may be in effect here that we might expect to hear melodies reverse their directions after big interval jumps (Margulis and Levine, 2006, p. 176). Because of this factor, we can exclude this example from a melodic point of view; and think it as a reference to define/re-define the melody when needed. Similar kinds of illusions can be found in many examples in -especially the recent part of- music history. In Figure 2.8, Deutsch (2013) gives an example from the final movement of Tchaikovsky's 6th Symphony while she explains "the scale illusion"; the upper staves represent the actual lines that are performed by violin groups, where the lower ones represent the perception of it.

This example can be seen a musical application of the same principle of grouping when negated. The illusion can be possible with one instrument or two, or more (groups of) same type of instruments; but not on different type of instruments with very distinctive timbres. Deutsch asks the question of "how does timbral difference affect the illusion" and gives several examples including a study of Gregory (1994) where he constructed various combinations of ascending and descending scales with different instrument timbres (Deutsch, 2013, p. 219). His experiment showed that amount of difference in timbral qualities affected the perception of the lines, and timbre became the main reference for streaming.

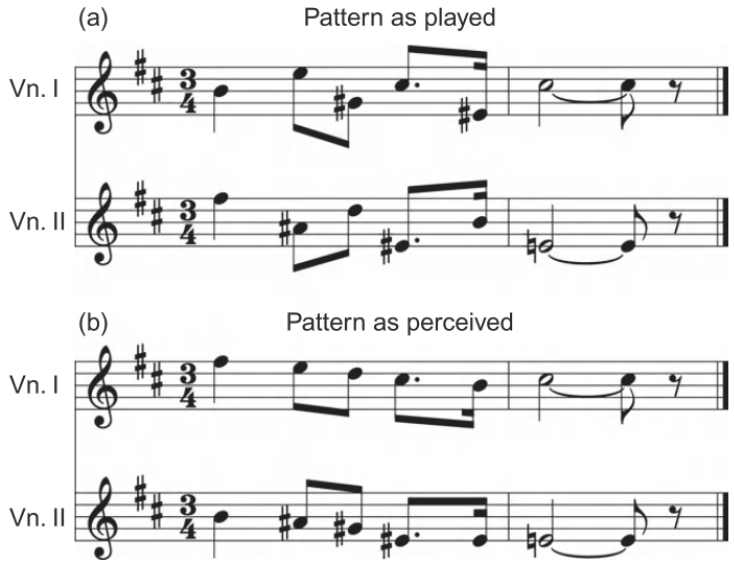

Figure 2.8 : Beginning of the final movement of Tchaikovsky's Sixth Symphony (*Pathétique*) Deutsch, 2013).

A more detailed study dated by Iverson (1995) focused on the topic of audio stream segregation. In his experiments, he included several acoustic attributes of the timbre. In the first experiment, participants were given repeating sequences (with no pitch variation) composed of two instrument sounds and they are asked to rate how strongly instruments segregated. The second experiment was focused on interleaved melodies and participants tried to recognize the melodies by a target timbre.

Iverson's experiments had three main findings: (i) Streaming judgments of participants correlated similarity judgments. (ii) Tones with dissimilar spectra segregated better. (iii) Dynamic attributes also influenced streaming. Streaming was

influenced by sum and difference of PATs[37] (timbres with shorter PATs and timbres with bigger differences segregated better), and the sum of spectral flux (less spectral change caused better segregation). The results from this study showed that both static and dynamic attributes of timbre can influence stream segregation and streaming was influenced both in presence and absence of pitch variation.

This part of the melody-timbre relations could seem relatively clear; yet, it is important to remember these results when a consistent musical grammar is desired by the composer. Every streaming whole creates another layer of implications in musical context.

2.3.5 Expectation and memory

The final stage of perception of timbre is related to memory. We have representations of sounds in our memory and they are stored like any other information. Unlike visual perception, which is distributed over time, an auditory moment is usually not repeated in time; unless it is a characteristic part of a music, or a sound object. At this point, echoic memory[38] helps us remember sound for a short time. Through memory, the sound analyzed by auditory system is linked with our previously categorized and stored references in long term memory. If the new stimulus is similar enough to any previously categorized representations, it will be considered as a member of the same category (Pressnitzer and McAdams, 2000, p. 55).

The basis of this stage is mostly at the level of attention and, at the cultural familiarity of the sound object that the listener is exposed to. Our memory is shaped by experiences, while expectation is based on our memory and the systematic re-interpretations of them. Consistent with the recent rising ideas in memetics, some recent studies show that memorability is very much related with our expectations that are rooted in our cultural routine and elements with expectancy violations are more memorable, thus, they have higher chance to be propagated (Porubanova et al, 2014).

---

[37] PAT (Perceptual Attack Time) is a subjective measure of the time that a musical sound's rhythmic emphasis is heard.

[38] Echoic memory is auditory the equivalent of visual “iconic memory”, and it is essentially the same as what has been called “pre-categorical acoustic memory”. It has been used as a synonym for short-term memory, however, while short term memory can last up to 30 seconds, echoic memory is limited to half a second (Winn, 2001).

In music, this relation is essential if we want to understand the salience and balance of structural elements.

The notion of expectation is very present in goal oriented musical grammars. In tonal music, harmonic and melodic structures are shaped together towards the same goal. In this combination, harmonic progression and instrumentation shapes the timbral properties of music; thus, timbral qualities contribute to expectation through tension (Paraskeva and McAdams, 1997). Reverse is also true; perception of timbre can be influenced by melodic properties.

In a study by Margulis and Levine (2006), (musically trained and untrained) test participants were given two timbres and four melodies. They were asked to identify the timbres, first in isolation, then as last notes (probes) in these melodies with alternative pitches. These alternative pitches were also ranked in their tonal stability and interval proximity to the previous notes.

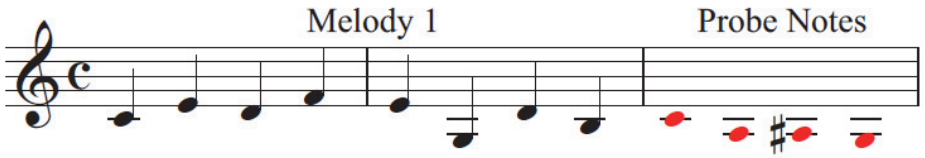

Figure 2.9 : A melody and its probe notes, adapted from (Margulis and Levine, 2006).

First important result from this experiment was that timbre identification "sometimes" improved when the timbre was identified when appended to a melody, rather than in isolation, which means melodic context helped timbre to be identified easier. Secondly, this improvement only occurred when the final notes had congruent (in the tonal context) pitches. When the pitch was incongruent, identification was not affected or got worse. Third, when there was confusion about the tonal context, stability overran the proximity factor.

This tendency in listener's perception shows that expectancy improves the perception of timbre. It is also an indication that melodic expectations facilitate timbre identification and (consistent with the nature of tonality) these "expectations are formed at a relatively early level of processing (Margulis and Levine, 2006, p. 181).

Lastly, results show that musically trained and untrained were both active in listening, meaning, melodic expectation did not require much training.

However, abovementioned study used melody only in one context and it is essential to emphasize here that the researchers do not use the term “tonal melody” in the paper, and they use “tonal context” only once. So, the term “melody” is directly associated with tonality without any explanation given beforehand; and it may still be question if timbre identification can improve with other types of melodic contexts. This can be seen as an indication that it is a matter of musical culture and individual’s exposure to it.

A similar conclusion was made by Wolpert (1990) in her study that “examined the salience of instrumentation over melody and harmonic accompaniment in identifying music excerpts”. She studied both with musicians and nonmusicians and used a method that tested the memory with changing musical data. The main question was about the perceptual dominance of melody and timbre in participants with different musical profiles. Results showed that nonmusicians showed greater response to instrumental timbre change, while musicians were more responsive to melody and harmony. Interestingly, musicians’ performances in these tasks were 100%. The strong conclusion of this study was that nonmusicians did not share the same cognitive schemata as musicians (or trained listeners) do; and she added that melody and harmonic accompaniment were the essentials of the Western music, “while the instrumentation was cosmetic” (Wolpert, 1990, p. 95).

Another study by Radvansky et al (1995) questioned Wolpert’s conclusion. Main criticism was that the instructions given to the participants were vague in Wolpert’s tasks. They did not specify a basis for recognition, listeners were told to select options with the “most resemblance”. This was true, but at the same time it was intentional; because “it was important that the instructions not direct the subject's attention to any one aspect of the music, nor in any way make clear the musical assumptions of the culture (Wolpert, 1990, p. 102)”. According to Radvansky et al (1995), this intention and the conclusion was not entirely consistent; Wolpert’s data showed different interpretations of the instructions and not the differences in mental representations. Their first experiment included nonmusicians and they were given both ambiguous and unambiguous instructions in separate groups. Results show that nonmusicians mostly used timbre in making their memory judgments. They made less error when

given unambiguous instructions and both groups showed the same error pattern, "suggesting that timbre is relied on in making memory judgments and is a substantial component of memory for music, at least for nonmusicians" (Radvansky et al, 1995, p. 135). Their second experiment included both musicians and nonmusicians; and used the same data and setting as the first. Results showed that nonmusicians performed worse on music memory task. But a more remarkable result was that musicians were also influenced by timbre changes, just as nonmusicians. So, they concluded that despite their difference in memory for melodies, there was no reason to link this to the reliance on timbre as Wolpert suggested.

The evaluation of these two studies brings us back the 'cultural' profiles. There might be a logical fallacy in motivation and conclusions of Radvansky et al. When their experiments used direct instructions for nonmusicians, (recognizing that there are no differences in mental representations) they basically tested 'how musician' were the nonmusicians. From this perspective, they have possibly underestimated the importance of the main result of Wolpert's study, which showed 'how nonmusicians perceive the musical hierarchically', at least in that certain culture.

Nonetheless, both articles basically recognized timbre as an "integral part of a novel melody" which helps to identify it later. A later study by Lange and Czernochowski (2013) investigated the same topic but their experiments did not include professional musicians, and involvement of some participants was only at recreational level. They examined the effects of timbre change on familiarity and recollection using old and new melodies. Their results showed that timbre change affected the memory regardless of individual's conscious recollection or familiarity. The findings of the study supported the idea that "timbre is stored in memory both as a context cue and as part of an integrated representation" (Lange and Czernochowski, 2013, p. 143). In short, previous studies strengthened the idea of melody-timbre integration in memory. Especially the difference between musicians' and nonmusicians' perception of timbre shows the role of it in identification of musical structures.

2.4 A Re-Evaluation of Timbre

An inclusive map of timbral relations needs a certain hierarchy and categorization in different aspects. The similarities and differences between timbral musical structures need to be clarified. Therefore, the topics already discussed throughout this chapter

make serve as basics to devise the possible hierarchical categorizations that will be presented in following sections; and these will be helpful to develop the totality of the proposal in later chapters.

2.4.1 Timbral levels

In his 1987 article "Timbral Hierarchies", Lerdahl writes the following sentence: "The real reason, I would argue, why timbre has been regarded as a secondary musical dimension is that, unlike pitch and rhythm, it has lacked any substantial hierarchical organization." (Lerdahl, 1987 p. 138) Then in the same article, he continues his search for possible timbral hierarchies through topics of grouping, prolongation, timbral intervals, timbral consonance and dissonance. This is followed by creating a 'vowel array' functionally similar to a proposed timbral classification. Towards the conclusion, he says that "timbral hierarchies can be developed much further, but that the task will be arduous if the goal is — as it must be — musically beautiful sounds." (Lerdahl, 1987, p. 155)

Especially aforementioned studies of McAdams and others that are done after Lerdahl's proposition, negate many assumptions in it. Depending on the previous findings, it is necessary to point out these considerations:

- "The real reason" might be a strong argument, but it is clear that timbre has a set of advantages and disadvantages when compared to pitch and rhythm. As stated before, a timbre is not recognizable until 200 ms of duration and we cannot put different timbres in order if they are under 500 ms. This is a disadvantage if a certain musical culture values rapid temporal elaboration, where timbre naturally becomes a secondary aspect.[39] However, in some other musical cultures that has varying levels of tempo and rhythmic composition, timbre can be primary concern. [40]
- Timbral intervals exist only in physical analysis and when the perception is directed towards certain parameters (See 2.3.3).

[39] In cases of re-listening and directed focus, even in non-timbre focused musics, timbre becomes the concern sooner or later.
[40] A notable example of this can be Pansori tradition of Korea. Use of human voice and descriptions used for them in this tradition can be seen in (Willoughby, 2000).

- Timbral consonance and dissonance (as described on the terms of stability, harmonicity vs. inharmonicity, in Lerdahl's article) cannot really be functional in a way that it functions in tonal music. Stability of a timbre based harmonicity only promotes tonal hierarchy in a traditional sense. Moreover, this idea is politically incorrect in the light of some recent studies that demonstrate concepts of consonant and dissonant are not innate; and some peoples of the world find "consonant and dissonant chords and vocal harmonies as equally pleasant." (McDermott et al, 2016, 547)
- Lastly, prolongation of timbral structures is of course possible, like any other musical element (e.g., gestures, see Chapter 4). However, this is directly related to the recognizability of the given timbral object. In fact, 'absoluteness of the timbre' makes it easier, and groupings through timbral similarities can thus be possible.

Therefore, categorization of timbre cannot be based on conventional idea of musical hierarchies. Timbral types of different qualities can have perceptual advantages over some others; however, a generalization of this kind of relations in common musical timbres is almost impossible. Instead, firstly a description of timbral levels can be done in order to indicate the scope of the timbre that is being described. I propose the following four levels for timbre:

- Microtimbre: The most basic level of timbre. A defining unit of timbre coming from any source; and does not change its basic perceptual qualities in any case; a 'timbral unit'.[41] In an instrumental context, microtimbre is a timbral (thus, registral) region of an instrument that has characteristic features that distinguish itself from other regions.[42] Microtimbre is naturally open to modification, however, when it is modified, it has to be labelled as a new timbral unit.
- Macrotimbre: This term used by Sandell (1998) to define the total timbral features of an instrument for all timbral regions (as cited in Siedenburg et al, 2019, p. 152). Therefore, macrotimbre can be seen as a sum of all timbres

[41] Bausch (2018) uses the term 'spectral atom' for a very similar concept.
[42] E.g., Clarinet registers *chalumeau*, *clarion*, and *altissimo*.

created by an instrument. The level of congruency along the axes of macrotimbre, points to a basic level of timbral coherence.

- Overall timbre (compound, group, ensemble, orchestra): Overall timbre marks the second level, (mostly, short term) timbral features of any musical ensemble in vertical axis. Here, vertical axis should be emphasized; because when a timbral region with a longer duration has a significant change in timbral components, it cannot be considered as the same timbre anymore. Thus, by definition, various overall timbres follow each other in a given work. Every regional relation becomes the constituent of timbral coherence. At this level, the equivalent of timbral units can be named as 'timbral objects', in varying magnitudes, durations and densities. (A timbral unit is still a timbral object, the opposite is not always valid) These timbral objects are not just combinations of spectra, but a new totality, a new gesture in timbre. François (1990) defines this term and extends the notion to "the whole piece" (p. 115). For the sake of levelization, here the totality of a work will be used separately.
- Formal timbre: This is the totality of the timbres that are presented during the whole work. However, this level marks two different things. Firstly, it is the level where we can see a formal scheme of timbral regions. Secondly, the sum of these properties also indicates a general characteristic for the work, the timbral totality, the atmosphere, 'the sound'. Therefore, basically, formal timbre to overall timbre in temporal axis can be seen as analogous to what macrotimbre is to microtimbre in vertical (spectral) axis. All timbral levels contribute to formal timbre, regardless of context.

These levels can be grouped into different sets. Firstly, microtimbre or 'timbral unit' can be seen as the main definition of first level. This unit cannot be modified into its components without losing its main properties. Macrotimbre of an instrument (a physical, traditional instrument[43]) becomes a combination of similar unit timbres, is an extension of the first level, despite the fact that it has more connections with temporal axis due to its realization process. Thus, it can be considered as in both levels depending on the context.

[43] Electronic instruments and different instruments combined in sets are excluded from this definition. However, it is still possible to extend the meaning of macrotimbre, taking the instrument as a performative unity.

Second-level timbre definitions include the overall timbre and formal timbre. Consistent with the nature of timbre cognition, the transformation of the overall timbre shapes timbral sections, and finally formal timbre. In cases of solo instrumental works, instantaneous microtimbre reaches the level of overall timbre; and expansion of it over the temporal axis shapes the formal timbre. It may also be useful to think that the notion of 'texture' is realized between the borders of overall and formal timbres; and this happens through regional timbral properties and gestural organization.

These level-based relations are summarized in Figure 2.10.

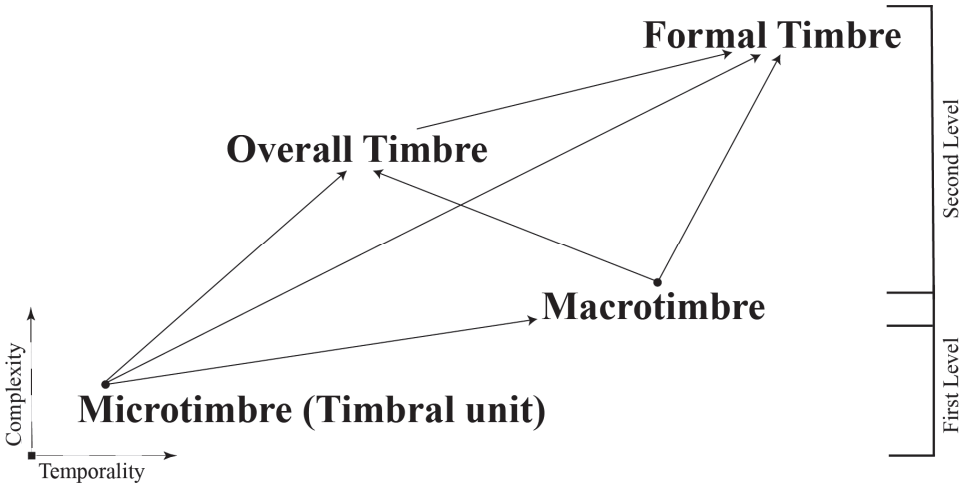

Figure 2.10 : Timbral levels and their relations.

This setting of timbral levels, alongside gestural organization, will be reconsidered in Chapter 4.

2.4.2 Timbral types

There are two main approaches in the subject of timbre analysis. One of them is the positivist approach that takes the timbre as a physical reality and quantifies it using computer technology, visualize it using spectrograms and when needed re-synthesize to a certain level of definition. Second approach is the one that defines timbre through its cultural aspects. Both of these approaches had proponents and some recent work is done to combine them.

At this point, it is necessary to admit that this thesis is a discovery of second approach's cruciality. Any attempt to study and analyze timbre will fail without a fair consideration of its physical complexity, perceptual uniqueness, and cultural aspects.

Following type descriptions will include both cultural and physical inputs. They will be combined through inferences made in the previous sections.

The term ‘timbral type’ is one that almost reaches to be characterized as an oxymoron. The adjectives and categorizations used for different timbres are most of the time problematic. However, it is possible to generalize these type-based categorizations with taking our basic analytical descriptions into account. Following descriptors will be explained shortly and will be contextualized in Chapter 4.

- Source identity: Different sets ranging from solo instruments to small ensembles and to orchestra. The logic behind using this is to simply not over-complicate the proposal, because every instrument name is already a type in musical culture.
- Registral region: Physical instruments have certain ranges and any art music tradition uses these instruments ranges in different registers quite carefully, in order to get certain timbral results. Macrotimbral properties, including different timbral properties in every register, are one of the first topics in instrumentation education. It is the same for orchestra, since orchestra is too an instrument, but with more modifiers in every register.
- Spectral region: Overall spectrum of an ensemble is directly related to its instrumental and registral components. However, same registral properties may result in different spectra with changing balance in dynamics and modulations. Due to the nature of timbre, spectrum of any musical sound changes rapidly both in instrumental or ensemble level, and in micro-temporal scales. However, these changes rarely affect the perception of the overall timbre.
- Timbral Envelope: Timbral envelope can be examined in two levels and they are fairly similar in morphology. First of these levels belong to a timbral unit (microtimbre) and occurs as an instrumental timbre in a very short time. Second one can be examined in the timbral regions in any given music, regardless of the instrumental combination.
- Modulations: This term can be used simply as any modulation technique or condition applied over the basic sound producing method (as it is also going to be used in next chapter).

The re-evaluation of the covered topics lead to the level-based organization of timbre. These levels and the approach based on types and classes of timbre will be the framework for later chapters.

3. GESTURE

Audiences have experienced a great rate of change in music listening practices over the past century. Radio, television, but especially recording and reproduction possibilities have enabled almost everyone to experience music in ways that were not possible before. As music has become more available, some of its essential components went through changes in their functions. Most of what we are listening to or being exposed to have no visual elements; however, with visual accompaniment, other layers of communication occur. As a result, these developments bring out the need to look further into the issue of gesture, which is rooted in physical and visual part of music.

Gesture is one of the essential topics in music; but as a term, it is most of the time not clearly defined or difficult to define depending on the context. It has frequent use in academic fields. Yet, it becomes even more difficult to define and correlate gestural notions within the compositional process that extends out from the composer to the listener. Main goal of this chapter is to establish a basic understanding of musical gestures and categorize them in a hierarchical scheme, and to evaluate them from analytical and compositional perspectives. In order to achieve this goal, the first meaning of gesture should be examined in evolutionary and linguistic connections.

3.1 Gesture and Language

Gesture can be thought as a second layer in language or an attempted language in absence of a lingua franca (Haviland, 2001, p. 83). Thus, gesture becomes a lingua franca itself.[44] Here we have to deal with ancillary (or accompanying, or accompanist) gestures, the gestures we use besides speech or any other activity. Ancillary gestures do not have a grammar, they have no additional layer of meaning, and they are holistic "just as a musical phrase is holistic" (Mithen, 2007, p. 154). On the other hand,

[44] The gestural similarities between neighboring cultures with completely different languages can be an example to this.

intentional gestures are symbolic (or emblematic, like todays language or writing system-based gestures). Same way, we can think of pre-language gestures as 'iconic', just as ape gestures, they show the path to do something, they manipulate, and they are not completely informative.

Since it is mostly rooted in cultures of humans, gesture needs to be examined in a larger context. However, following the same logic, we can also mention that birds have specific (body) gestures that accompany singing (Salwiczek and Wickler, 2004, p. 167) regardless of whether it is learned or genetically programmed. Same way, singing and non-singing primates have some similar gestures as humans, accompanying their songs or calls, such as branch-waving and beating of the ground in groups. Merker (2000) also states that this kind of behaviors "were the root of musical behaviors in our hominid ancestor" (p. 8). It is necessary to note here that, animals with beat perception and synchronization (BPS) (i.e., entrainment[45]) are mostly capable of vocal mimicry behavior (Morley, 2013, p. 245); and this 'mimicry capability' is the key to musical and linguistic development.

Gesture is also an important component of speech. Through some tests (i.e., with stutterers) we know that they are not autonomous (Morley, 2013, p. 229). Moreover, gestures have an evolutionary priority over verbal tasks.[46] Gestural combinations with vocal (musical) behavior could have led to more complex vocal systems. In addition, it is a remarkable fact that first children to produce gesture-word combinations are the first to produce two-word combinations (Clark, 2009, p. 156). Besides developmental logic, Patel (2003, p. 674) stated that although studies have shown that music and language capacities can be partially or entirely disassociated, the brain imaging studies suggest that they share the same neural networks. Language and music can share the same analytical system, in the end, they have similar (actually same and a deeper level) auditory properties. However, we use these data (which is fairly similar) for quite different functions; furthermore, we categorize them very differently. Despite they are produced by same mechanism, we 'think' they are different. As it seems, from an

[45] Entrainment: Synchronization of a biological rhythm, defined as phase and period control of an endogenously generated biological rhythm by an external cycle (Winn, 2001, p. 561).
[46] Morley (2013) gives a reference to Feyereisen's (1997) study that showed in a situation where manual and verbal task was given, that manual task had no delay while verbal task could be delayed, thus showing that gestures have an evolutionary priority.

evolutionary perspective, they evolved together "with neither modality taking precedence over the other" (Kendon, 2016, p. 163).

From a more practical perspective, language is music depending on the context. As Ihde states, "the foreign tongue is first a kind of music before it becomes a language" (Ihde, 2007, p. 157). On the other hand, when there is a meaning perceived, the sound component of the language stays behind.

> Language-as-word, unlike music, even while sounding, does not draw attention to itself *as* sound. And yet, were the other to be speaking, and suddenly the sound actually disappear, I should no longer be able to hear what was being said. (Ihde, 2007, p. 158)

The focus changes from gesture to semantic implications of phonetic combinations as the music and language is separated; and the semantic units are gradually strengthened and linked to another mechanism. However, it is necessary to remember that the phonetic units are essentially 'signature timbres' that also lose their sonic features with the inclusion of semantics, it is not only the quasi-melodic content of gestural contours.

To conclude, basic level examination of this relation presents us three points to consider in the examination of musical gestures: (i) Gesture is one of the basic building blocks, both in verbal communication and music; and they are shaped in our cognitive mechanism starting from our birth. Thus, musical equivalents of gestural roles and functions could be essential to comprehend the overall mechanisms and structures in music. (ii) Some gestures can be holistic, as in ancillary gestures, may not imply anything other than the effort put into a certain action; some gestures can create an extra layer of meaning, they can be 'symbolic' and can have a function of directing the focus (as will be explained in communicative gestures). (iii) It is important to consider the balance of gestural and timbral contents in music, as they may have a similar 'hiding effect' for each other, and in their own networks.

3.2 Musical Gestures and Their Categorization

Musical gesture begins with understanding the gesture in its defining character. As Jackendoff and Lerdahl (2006) state, "the term 'musical gesture' is directly motivated by its relation to physical gesture" (p. 65). Physical gestures form a layer of intensity and confirmation in communication. When a musician plays an instrument, an energy finds its way of flowing out through gestures particular to that certain practice and

culture. A guitarist is expected to show guitaristic gestures and a clarinetist is expected to move the bell of the clarinet up and down; although it does nothing to help sound production (Vines, Wanderley, Krumhansl, Nuzzo & Levitin, 2003). The movement of the instrument is an extension of the musician's physical gestures (if it is a movable one, of course) and the body, becomes a part of the instrument itself. Thus, the mental part of the musical gesture is not only related to cultural motives and identity, but also to body gestures, and ergonomics of playing an instrument.

Hatten (2004) states that "gestures are not only actions to create sounds, 'but the characteristic shaping that give those sounds expressive meaning" (p. 93). Because as they are used more often, in time (short term or long term), they tend to create an additional channel of communication. As he states, gestural meaning emerges from "symbolic motivations based on enculturation in a musical style" (Hatten, 2004, p. 94); and it is necessary to remember this usage of 'meaning' can only point to an implication or an indirect reference, not a real meaning. Despite the necessity to question the concept of 'meaning' here, this is quite consistent with the idea of the gesture as 'the energy that shapes' and it is useful to see the notion from this perspective when discussing its compositional use.

3.2.1 Performer's gestures

Performer is the principal interpreter and conveyor of composer's gestural complex on stage. Therefore, the gestures performer uses are to be categorized first, in order to see the relation between compositional intentions and their active representations. Performer's gestures are classified, by Delalande (1988) as (a) effective gestures, (b) accompanist gestures and (c) figurative gestures. Effective gestures and accompanist gestures can also be called referred to as instrumental and ancillary gestures, respectively (as cited in Wanderley, Vines, Middleton, Mckay & Hatch, 2005, p. 97).

Effective gestures are the physical movements that produce the sound, regardless of what the visual result may be. This definition may seem to be in a direct contradiction with the main definition of gesture, because this physical movement is the effort that creates the labor; there is no saved energy and it does not 'mean' anything or imply anything else. The effort 'is' the reason why a performer has other types of gestures. However, the instrumental gesture can be accepted as a gesture if it is coming from a mental level, by defining 'how' the effort is applied. The selected sound producing

method is already about a gestural preference related to the nature of the desired sound; thus, the first link between the gesture and composition can be found at this stage: instrumental gesture has the mental part as the source and the physical part as the outcome.

Cadoz's instrumental gesture typology (Cadoz, 1988, p. 7) clarifies the distinction between the physical and mental parts of the instrumental gesture. He sums up these gestures in three types: (a) excitation, (b) modification and (c) selection. “Excitation gesture” is the action that applies the energy to the instrument. It can either be instantaneous (i.e., percussive or plucking) or continuous (i.e. bowing or blowing). “Modification gesture” is the action that creates small variations in sound. It can either be parametric (i.e. vibrato, sul tasto etc.) or structural (e.g., using a mute, applying the pedal functions of the piano, but not pressing the pedal itself). The last one in Cadoz’s typology is the “selection gesture”, which can be defined as the selection of multiple elements of the instrument sequentially or simultaneously. This gesture type is essential to the production of melodic and harmonic structures by selecting certain keys, strings or similar actions; however, it can also be considered a container of every type of instrumental gesture when they are combined in order to create a desired compositional output.

“Ancillary gestures” (or accompanist gestures) are the body movements (including facial expressions) that are not necessary to produce the sound, but which somehow provide a layer of confirmations to the music. These may be rooted in a certain musical culture or the individual's conscious and unconscious collection of gestures from various sources. Some of the ancillary gestures can also be used indirectly to modify the sound or affect the perception of the sound. This can either be achieved through exaggeration of the physical gesture, creating a certain expectation about a parameter in music (e.g., raising the hand more than necessary before the excitation to emphasize the moment of an impact [on a guitar or piano]; or using a bigger portion of the arm and a faster movement [without changing dynamics] to emphasize a staccato on any instrument).

All gestures are communicative on different levels. On the other hand, “communicative gestures” (or more precisely, “direct communicative gestures”) can be described as a specific type and a shared subset of the above categories. They are gestures of any other type or a new bodily (or sometimes instrumental) gesture that is

'intentionally' used to convey a message which is integral to certain aspects of the performance. As mentioned before, they are mostly 'iconic' gestures. This type is especially common when communicating with other performers in order to conduct a unified performance, in aspects such as rhythmic synchronization, dynamic balance or any expressive output. This can be either between members of an ensemble, or between a conductor and performers. They are physical gestures that convey direct, literal messages which are selected from a repertoire of gestural codes; and they work through a mechanism of directions, warnings, and confirmations. Communicative gestures (together with ancillary gestures) are to be seen through their performative functions. If they exist in compositional stage to serve a specific purpose, they are mostly theatrical elements; and their relations should be examined depending on the genre; and they are lost, when the music is recorded.

In Western Art Music, bodily gestures as visual cues are the most used; however, in some musical practices, direct musical gestures are also used without any concern of aesthetic discord (e.g., in a rock band, drummer giving the tempo, timing or the meter through a "count off", by using any parts of the drum set, and sometimes any member using the phrase "one-two-three-four").

A second type of communicative gesture can be described as between musicians and the audience, and most of these can also be placed in the category of ancillary gestures. In this case the semantic aspect of gestures may be partly ambiguous. Aforementioned facial expressions fit in this type when they do not occur as reflections of the performer's challenge during the performance, but as a channel to convey a personal (or collective) emotional significance (even if the emotions are not genuine at that moment). They are used to create an extra layer of expression, support a narrative; and they are aimed to affect the perception of the audience through their mood. Besides ancillary gestures, many verbal or bodily gestures are used to communicate on stage in popular genres. This is mostly restricted to the composer's intentions and directives and the 'fourth wall'[47] is rarely 'broken' in Western Art Music practice.

[47] 'Fourth Wall' is a term mostly used in theater, cinema, or literature; indicating an imaginary wall that separates the parts; it is 'broken' when performers interact with the audience or address them. It can also be applied to any staged musical performance.

Lastly, “figurative gestures” are the “sonic gestures perceived by the audience by means of the produced sound that have no direct correspondence to physical movement”, as described by Delalande (as cited in Wanderley et al, 2005, p. 97). This type can also be called a ‘received mental gesture’, because it is the result of a series of gestures transferred beginning with the intended mental gesture, the idea that is formed in the composer’s mind. Intended and received forms of a certain mental gesture will be different to varying extents, depending on the individual’s gestural codes, and listening conditions

3.2.2 First-level gestures

As the types and definitions of gestures proliferate, it is necessary to constantly mark the level we are on. The two main labels we used for gestures so far are ‘physical’ and ‘mental’. Another labeling, or rather, categorization I propose is level based: ‘first level’ and ‘second level’. This hierarchical categorization is both about levels of immediateness, and significance on a perceptual level. Nevertheless, these two factors do not always have congruency in a specific direction within this hierarchy.

First-level musical gestures are either directly related to physical gestures or they have similar properties without having explicit physical connotations. Hatten (2004) describes gesture as “a unit in ‘perceptual present’” (p. 94)[48]. There are time limits for gestures and on this level, these ‘unit gestures’ are analogous to the prosodic units[49] in speech. This is related to the idea of “expressive unit gestures”[50], a key term to grasp the first-level definitions of gesture. Ben-Tal (2012) suggests “expressive unit gesture” as an analogue to physical gesture, a term for the main definition of gesture in music. He points at the basic and important features of physical gestures like their durations and intelligibility. Physical gestures are short, they have clean contextual intentions and “emotional or conceptual signification”, and most importantly are “self-contained" (p. 251). A self-contained gesture can be a combination of several parts, smaller structures, but these structures must be in a balance as to make the gesture be

[48] “[Perception] segments experience into a succession of units whose relative autonomy is a function of the coherence of the perceptual information and perceptual strategy observer. Each of these durational units is a 'perceptual present'” (Clarke, 1987, p. 229).
[49] Prosodic unit (or intonation unit) is a term in linguistics and it is the smallest unit that combines a pitch contour, rhythm, and phonemes.
[50] A structural and functional analogy can be drawn between ‘timbral unit’ and ‘unit gesture’.

perceived as a total body that is not reducible to its elements. As a result, expressiveness of a unit gesture does not depend on the context; it can still be expressive independently.

From a compositional perspective, a possible confusion about the mental gestures would be the differences between gesture, figure, and motif. Comparing gesture with these two concepts are essential since they are two building blocks in the traditional craft of composition. There is a general tendency to use motif and figure interchangeably, or in the same category.[51] However, this approach excludes the functional features of these structural elements. A figure can be defined[52] as a short, non-expressive unit; it is mostly a basic cell of an ornamental melodic sequence, a short rhythmic unit, or a simple harmonic progression; an infinitesimal unit open to repetition without being a foreground element. On the other hand, a motif is larger and has the capacity to be the theme and core of a composition. However, in some cases of minimalist compositions, figures can also replace other elements and the composer "invites us to hear only background" (Scruton, 1999, p. 63). [53]

A gesture can also be used to create a thematic movement and any gesture can be used thematically, depending on the context. As Hatten (2004) states, "Certain motive-length gestures may be marked as thematic for a movement, hence foregrounded and amenable to development, variation, or ongoing evolution by means of developing variation" (p. 94). However, if a gesture is distorted or developed in certain ways, it may lose its initial properties as a gesture. To exemplify, we can see that the main motif used by Beethoven in his Fifth Symphony is a gesture in itself, as it is even given[54] an extra-musical connotation of 'fate knocking the door' (Schindler, 1996, p. 147).

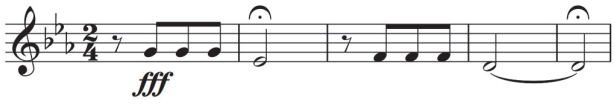

Figure 3.1 : Main motif from Beethoven's Symphony No. 5, Op 67, mm. 1-5.

[51] "A short melodic idea having a particular identity of rhythm and contour, often used repetitively or in conjunction with other such ideas to build a larger melodic idea or a theme. Thus it belongs to the category of musical ideas commonly called motifs (see Motif)" (Grove Music Online).
[52] Above definition is consistent with Ben-Tal's (2012) and Scruton's (1997) views, however, it belongs to me.
[53] In this case, Scruton refers to Philip Glass's music.
[54] Not stated by the composer directly in any source.

However, this musical material is repeated and modified during the movement loses its initial features. Later brought back in exact repetitions during emphasized moments.

A good example of these three notions together can be seen in the first measures of Chopin's F minor Nocturne, Op. 55, No 1.

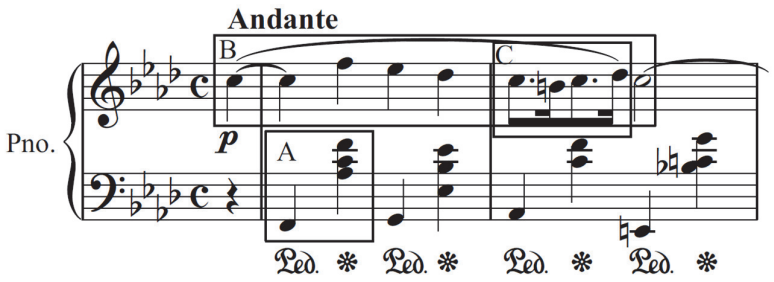

Figure 3.2 : Frédéric Chopin's Nocturne in F minor, Op. 55, No. 1, mm. 1-2.

In the score, given (A) portion is a figure, on the basis that its main function is to establish the texture through harmonic progression, whereas, (B) is a motif that is repeated and melodically developed further. However, (C) is essentially a gesture, it is not developed and sometimes disappears from the motif. Basically, motif is defined through certain functions related to other structures; whereas, gesture is an object, and it does not need to be functional in a systematic way. A comparison between these concepts are summarized as in Table 3.1.

Table 3.1 : Comparison between figure, motif and gesture, adapted and extended from (Ben-Tal, 2012).[55]

Figure	Motif	Gesture
Background	*Foreground*	*Foreground*
Regular repetition	*Repeated*	*Can be repeated*
Can vary in some dimensions	*Developed through composition*	*Salient characteristics preserved when repeated*
Incomplete on its own	*Can be complete but open to extension*	*Complete, extension counters gestural identity*
Can be notated precisely, result depends on instrumentation	Can be notated precisely, notation is the space of the motif	Notation represents the 'silhouette' of the gesture, the rest comes from the cultural connotations of the given material

[55] First four rows are taken or adapted from the cited article.

Table 3.1 (continued) : Comparison between figure, motif and gesture.

Figure	Motif	Gesture
Does not need interpretation	Does not need interpretation	Depends on interpreter
Not crucial in formal structure	Has the potential to create formal structures	Although it is not directly mentioned in traditional understanding of form, thematized or not, they can be important elements of form. (As it is discussed in 3.2.3)

3.2.3 Second-level gestures

Many similarities between physical and musical gestures help us understand their relation better. On the other hand, some of the differences are essential to explore their hierarchical positions in a larger network of structures.

Zbikowski (2011) states one of these similarities as the relation to language; and claims that unlike physical gestures that accompany speech, musical gestures are independent of language (p. 84). However, this claim excludes cases where physical and musical gestures can both be linked to language when there is language involved; in which, musical works with verbal aspects can share to some extent ("word painting"[56] technique is a good example of this). Nevertheless, it is not necessary to take language as a higher source of meaning here; even without any verbal aspect it is possible for music to have functional processes on its own. A second difference he states is as follows:

> While it is generally recognized that the gestures that accompany speech do not have a grammar, I would like to propose that sonic analogues are basic to musical grammar. From this perspective, then, the notion of a musical gesture is not a metaphorical one but a reflection of the essential materials of musical expression. (Zbikowski, 2011, p. 84)

Thus, if musical gesture is a part of a grammar (or may be the keystone to grammar in some cases), this grammar can be one of the skeletal structures in music. Through

[56] "The use of musical gesture(s) in a work with an actual or implied text to reflect, often pictorially, the literal or figurative meaning of a word or phrase. A common example is a falling line for '*descendit de caelis*' ('He came down from heaven'). The term is more usually applied to vocal music, although a programmatic instrumental piece might in some sense exploit the technique" (Grove Music Online).

repetition, transformation, and prolongation, they can build larger structures in temporal dimension.

Cone (1974) calls symbolic gestures (what is called 'mental gestures' here) "meaningless but meaningful", like exclamation words in language that function as "pure gestures" (p. 163). He sees musical gesture as a simulation of the gestural aspect in verbal utterance; and through this idea he reaches a conclusion: "If music is a language at all, it is a language of gesture: of direct actions, of pauses, of startings and stoppings, of rises and falls, of tenseness and slackness, of accentuations" (Cone, 1974, p. 164). According to Cone, the physical gesture is an emulation of utterance, an action that tries to speak, and he links this idea with the musical (symbolic/mental) gesture as a symbolic utterance. Ben-Tal (2012, p. 250) relates Cone's ideas with Wagner's idea of musical gestures as analogue to visual ones; and interprets Cone's understanding of gestures not as musical units but rather as 'energies in music that create meaning' (which is consistent with the initial definition of gestures given above).[57] Despite the accuracy concerning Cone's understanding of gesture, this does not mean these 'energies' are not musical units. On the contrary, they can be considered as essential musical units, but they cannot be considered within the first-level definition of gesture. We can see this type as a second-level gesture. In order to concretize this concept, a categorization through musical terms will be necessary.

All second-level gestures are mental gestures. They are either momentary changes in certain aspects, some threshold moments; or larger changes and possible perceptual connections on the horizontal axis. Momentary changes of various aspects like harmony, density, timbre, texture, rhythmic properties, or any other aspect that is emphasized can be included in the 'vertical' category. 'Horizontal' gestures on the second level are the phrases, cycles, or broader connections on the temporal axis, created through memory and cultural familiarity. These types of gestures (which will be referred as gestural bodies) may have analogue properties to unit gestures, but their perception in that manner is not always guaranteed because of the balance that is dependent on the factors mentioned in first-level gestures. They are not as strong as

[57] The definition of the basic physical gesture given with in this work was related to "saved energy by using a language or playing music"; which means that at each level, it is related to an energy that is left behind due to imbalance between bodily (genetic) and cultural evolutions.

unit gestures, because their impact is redistributed over time; considering they are not in the perceptual present but rather sometimes several minutes, or hours long.

Another important factor for second-level gestures is the context. Not unlike unit gestures, their implications may change according to the context they exist in, but in any case, they can be expressive. Despite his recognition of the gesture as a form-shaping element in music, Cone (1974, p. 165) argues that expressive content of the musical gesture needs a context. It may be true that a musical gesture will have a different implication depending on the context. However, this idea is to some degree a contradiction. The condition of something being 'without context' should mean it purely exists at a mental level. Since a musical element can never be without context once it exists as a sound; even the silence is a context and there can be a musical 'meaningfulness' created in the mind by the perceiver. Furthermore, supposing a given gesture is pitch-based, one (especially a musician) would easily be conditioned to perceive it in an imaginary harmonic context. That is to say, a musical gesture creates its own context even in the absence of an apparent one. [58]

A summary of gesture categorization can be presented as in Figure 3.3. "Compositional gestures" are presented as a subtype of mental gestures and "Communicative gestures" are presented only as a subset of performers gestures; in order to avoid confusion due to their complex relations with all other types.

Hierarchy in this type of categorization is different in each direction. Immediateness, simplicity, and perceptibility are higher on the upper part (first-level gestures); and complexity and inclusiveness in the part below (second-level gestures). In addition, gestures change from having direct physical analogies towards having more mental aspects moving from left to right. Nevertheless, it is always possible for some subtypes to be related to multiple categories; and hereby presented hierarchical relations sometimes can be reversed or modified in certain conditions.

[58] This is more valid today, as we are more used to listening our soundscapes and experiencing musically related art dependent on them.

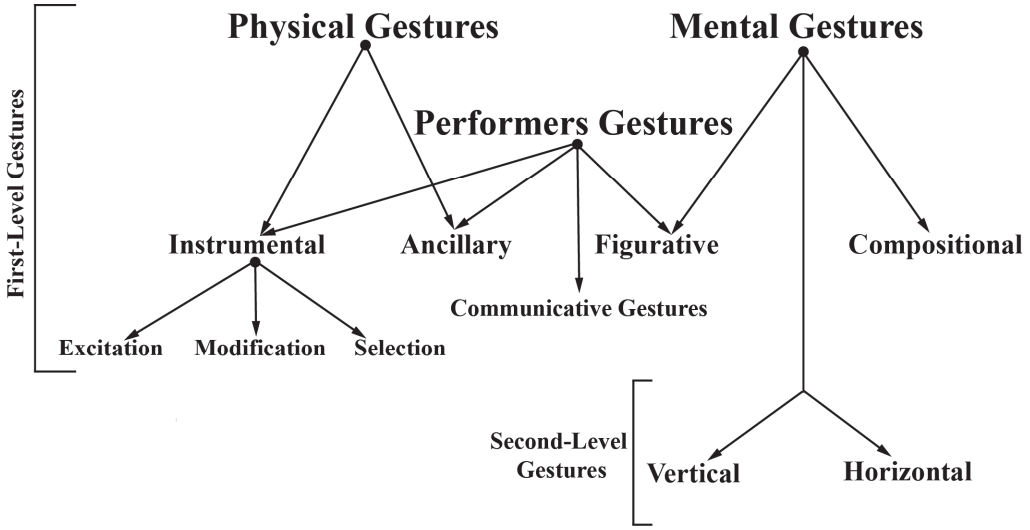

Figure 3.3 : Categorization of Gestures.

3.2.4 Composer's categorizations

Musical gesture's character is rooted in an identity. It can be specific to a certain culture, approach or a philosophical value of a composer or a performer. From a compositional perspective, they are the tools to utilize in various aspects. Therefore, a list of the possible compositional gestures may be necessary, or at least, useful. Composer's self-categorization and listing of gestures would help to evaluate them in larger contexts; yet, this list can never be complete.

Florence Baschet, a composer, presents a list of gesture types "based on a particular playing technique, to the collective musical gesture, or even compositional gesture" (Baschet, 2013, p. 22). These types are defined through a certain work, *"StreicherKreiss",* for string quartet and live electronics, which she composed as a result of her research at IRCAM[59]. She uses the concept of gesture in a wide range, including instrumental and compositional gestures together. Some of these types use very lucid notions; however, they can be used for more elaborate relations. For example, "silent gesture" is defined as a cello player acting like playing certain techniques but not producing any sound; and through these movements she/he shapes the sound of the quartet and electronics. Undoubtedly, this becomes a physical gesture; just like how the conductor shapes the orchestral energy through his movements, to

[59] IRCAM: *Institut de Recherche et Coordination Acoustique/Musique* (Institute for Research and Coordination in Acoustics/Music). French institution for music research. The institute is a home for avant-garde acoustic and electronic music (especially spectralism) since its foundation in 1977.

shape the orchestral sound.

As another example to the aforementioned issue of context (both in Baschet's and previously Cone's), Baschet (2013) lists "frozen gesture" and "nearly non-existent gesture"; and she adds that they may seem like oxymorons (p. 229). Actually, they are defined by musical expectation or liminality of perception. Scruton (1999) describes silence[60] itself in a similar context and says that "[silence] is not a cessation of action, but action of another kind – refraining, withholding, refusing. Silences in music are always pregnant" (p. 333). In this sense, extreme development, silence, or stasis of musical material can be included in the second-level definition of gesture.

One other type Baschet presents is the micro-gesture, which refers to various types of instrumental gestures with the simplest definitions (e.g. order of bowing directions, vibrato, certain changes in any parameter). When these micro gestures are used together, she calls them "the articulation of various combinations of playing techniques" (Baschet, 2013, p. 22). Because, this type is explained as a complex superposition of gestures, it becomes an independent, larger scale (a second-level) gesture itself.

Cases of implied gestural inflation create change in resolution of listening and, reduce focusing possibilities. Therefore, a need for hierarchical organization of gestures may arise. The hierarchy at this level is relative; because -depending on the features of a given musical sequence- unit gestures can be more on the surface than formal gestures, and vice versa. Thus, this balance can affect the perception of the form. This is applicable for both simultaneous gestures and gestures distributed over time. This is the second-level definition of gesture; here it is not a unit but a gesture with smaller gestures in it. Hatten (2004) refers to the larger gesture as 'formal', phrases, or melodic contours which are longer than the perceptual present and have the potential to shape the overall form. Finally, the balance between the macro and micro levels determines the gestural level in a moment of music.

In the context of gestural analysis, timbre is an essential concept. Baschet (2013, p. 23) defines 'timbral gesture' using the notion of uniqueness of the instrument and the player, and the impossibility of getting the same timbre twice. When thought in a larger

[60] Silence, without any evident effort to create sound. Unlike Baschet's "silent gesture" which is actually a physical gesture.

context, like she does with the complex of gestures above, this term can still be applicable because, the change and the control over timbre is a compositional gesture itself.[61] Furthermore, networks of instant and larger scale temporal transformations of timbre lead to formal structures; and this idea will be explained in following sections.

Lastly, another compositional approach to gesture is Philippe Leroux's use of notational calligraphy (Garcia, J., Leroux, P. & Bresson, J., 2014). Leroux takes neumatic notation of medieval music, particularly the manuscript of Guillaume de Machaut's palindromic rondeau *"Ma fin est mon commencement"*, as a source of gestures. On the technical side, neume units are captured by a special electronic pen and sent to pOM application ("pen-OpenMusic"[62]) to be interpreted in terms of contour, direction, speed of hand gestures; and this data is analyzed and converted to parameters like musical pitch, duration, rhythms, area, perimeter and some others. The result is used as compositional input for a new work. The idea of taking static representation of a music as a source brings forth the topic of how we perceive visuality in music related objects that are inherently dynamic. By activating the static representation, Leroux creates an augmentation of one of the steps in the conveyance of gestures, adds new layers to its transformation. Naturally, this transformation is not objective, there is a substantial involvement and subjective interpretation of the data. Nonetheless, the creative process is almost never free from them. In order to clarify these relations, the stages of gestural conveyance will be examined in the following section.

3.3 Conveyance of Musical Gesture through Composition and Performance

Musical gesture gains different aspects when it comes to using it in a compositional process. Firstly, the analogy between physical and musical gesture is not valid anymore; because, unlike instrumental and ancillary gestures, creation of the composition does not share the same space in time. The composer spends time thinking on it, designing, and modifying it; there is limited or no spontaneity. Secondly, it can never be directly conveyed from the composer's idea to the listener's experience. It has multiple stages for potential changes, realization, interpretation, and re-interpretation; until it becomes a figurative gesture in the listener's mind. Thirdly, composition is a

[61] This is one of the most common compositional concepts in early spectral music.
[62] OpenMusic (OM) is an object-oriented visual programming environment for composition.

process that has the potential to use almost all gestural types in it with certain relations depending on choices and coincidences.

I propose to analyze the conveyance of gesture in compositional process within the frame of following four stages:

1) Musical idea as a gesture, gesture as a musical idea (Any kind of mental gesture, compositional or instrumental): A musical idea may be dependent on a gesture; and composition may depend on the prolongation of that idea with the help of other structures. Composition may be constructed through prolonged gestures and every building block has the potential to be a gesture. As mentioned before, gesture should be self-contained and have a clear perceptual significance. (As an expressive unit gesture – or a clear second-level gesture)

2) Static representation of the musical idea (As notation, or in case of electroacoustic music, as an acoustic object): In this stage, gesture is partly abstracted from its energy. When it is represented on paper, musical gesture transforms into a representation; just as spoken language transferred to written form. Yet, musical notation contains more visually related data and an aesthetic aspect (See 3.2.4, example of Leroux). This gives the musical gesture another dimension to exists in. Furthermore, when musical gestures are translated into notation, they require a "knowledge of the relevant musical style and culture" (Hatten, 2004, p. 94). Notation represents the 'silhouette' of the gesture, the rest comes from the cultural connotations of the given material, instrumentation, and context.

3) Visual and acoustic representation of the idea (Instrumental gestures, ancillary gestures, accompanying visuals): During the performance of the music, previously thought and represented gesture arrives at a new level where it has a body of sound and occasionally a visual representation in the performer's body gestures. On this level, the acoustic reality brings the musical gesture to its final form before perception. In addition, lack of visible physical gestures in music (e.g., tape music, installations, electroacoustic music without significant physical gestures) make extra visual presentation functional as an accompaniment.

4) Gesture as a perceived idea (Figurative gestures): The final destination of the gesture is the listener's mind, where it is interpreted (or re-interpreted on re-listening) with all related data, including the performer or any other accompanying visual and cultural data from personal memory. Now the gesture has found its meaning, until it gets corroded or deformed by other culturally related data. Similar conditions as in notation are relevant when listening to a performance. It is possible that "gestures may be inferred from a musical performance even when we do not have visual access to the motions of the performer" (Hatten, 2004, p. 93). However, a broader consideration about the enculturation is necessary at this stage as well.

These four stages are visualized in Figure 3.4.

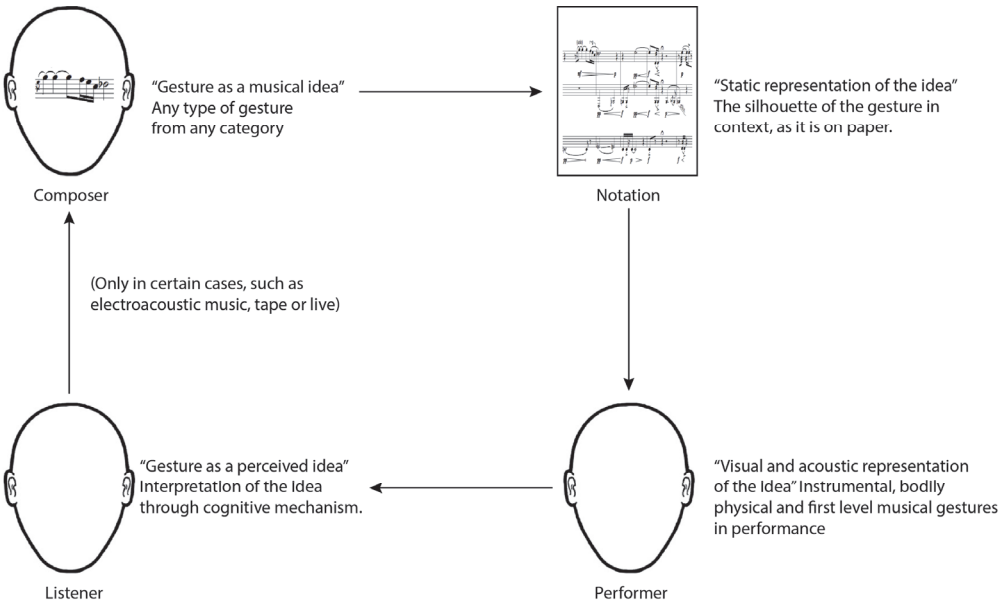

Figure 3.4 : Conveyance of Musical Gestures.

An additional comment about gesture in electroacoustic music would be appropriate at this point. In this case, stages can mix or disappear from the scheme.[63] Firstly, the production stage in electroacoustic composition may be different than the acoustic composition tradition. The composer can experience the work while composing and present it in its final form and it does not need a performer most of the time; and when

[63] An important reference here is Brian Eno's 1979 lecture titled "The Recording Studio as a Compositional Tool".

it is performed, physical gestures of the performer do not reflect the auditory counterpart; or in some cases with live electronic music performers, extra body gestures which are not essential for the performance are added intentionally. Thus, perception of gesture in electroacoustic work may largely be affected by the choice of auditory setup and visual presentation. On the other hand, when electroacoustic possibilities are used in combination with performers of acoustic instruments, gestures from different sources (including physical ones) have the potential to mix and create combined layers. In that case, gestural conveyance and its analysis requires settings specific to given work.

From this point on, identifying specific functional relations between musical gestures in combination with timbral structures is needed.

4. THE INTERACTION: GESTURE-TIMBRE NETWORK

The concept of 'objectivity' as mentioned in 1.1 does not imply Eurocentric[64] or similar claims of universality. The age we live in certainly is an age of multiple approaches, methods, existences, identities, cultures, and universes. However, regardless of the bias created by every single experience of these layers, the totality of life and our cognitive capabilities are bound with certain, highly similar, if not the same, principles. In addition, regardless of their semantic connotations that are open to interpretation in every other condition or culture, structures in music easily form greater networks. A cultural specificity that is failed to be perceived as intended, never fails to be linked to a network of structures when it is raised to a significant state, because it is a perceptual input, regardless of its intended meaning or function.

Timbre and gesture are not 'chosen', in this context. By definition, they represent the kinetic and quasi-static properties of all energies in music. Quasi-static here does not specifically indicate any of them, even if it is easy to relate it more with timbral features. However, it is reasonable to see them in a continuity of different levels of kinetic properties. In this context, in a simplified way, gesture can be the 'unit movement that represents the kinetic energy, and the energy that shapes musical structures' (See 3.2); and timbre as the 'distribution and movement of energy with certain balance over different timespans'. In a visual analogy, gesture is the shape (or a complex of shapes), timbre is the combination of colors that affect and affected by that shape. Thus, timbral object in a kinetic context is a form shaping gesture.

At this point, these notions will be served as network structuring forces that result into musical form. Timbral coherence, thus, does not represent a way of labeling or attributing a certain function; rather, a Gestalt of timbres and their transformation with several possible interpretations in other aspects. Similarly, the analysis of gestures and their networks do not have to 'mean' anything more than an invented language that

[64] Since the intended scope of the thesis is on the Western Art Music, it is necessary to point out that author of this thesis does not approach the subject from any cultural viewpoint.

sounds ‘musical’. As commonly experienced, the language one cannot understand is always more musical to the ears, and music does not have to mean anything. In summary, these principles avoid semantic analysis and concentrate in physical, physically definable attributes in both dimensions; and thus, they are not limited to analysis of a certain genre or tradition of music, but adaptable for potentially all.[65]

4.1 A General Description of The Method

The core idea of timbre/gesture-based analysis is similar to Schenkerian approach, where he says that the local harmonic structures can be summarized and related in groups in order to be combined in larger structures, and these structures can be simplified to a counterpoint of a descending melodic line and harmonic functions (Schenker, 1977, p. xxii, 3-5; Schachter, 1999, p. 186-187; Cadwallader and Gagné, 2011, p. 4). An analysis of form which is established through timbral and gestural layers would do the same; except it does not simplify to smaller structures, it does not “reduce the work to a single basic gesture” (Scruton, 2014, p. 239). Rather, it relates perceptually functional moments of intersections and changes to create the perceived form of the musical work.

The following proposal is a three-layer scheme for the analysis of perceived form:

> 1) Gestural layer: Unit gestures and larger gestures shape the perception and give defining characteristics on the surface. This layer consists of smaller gestural layers, the smallest being unit gestures; then second-level gestures and form-shaping gestures. This can be thought quite like the Schenkerian analysis of form: Foreground analysis examines the relations between smallest units, middle-ground analysis examines second-level gestures. However, the third level would be different than the background level of Schenkerian analysis; because, unlike *ursatz* (which is the skeleton of the work), gestural background cannot be an easy summary derived from the second level, but rather an analysis of interaction and perceptual hierarchy of the first two levels. The

[65] E.g., a tomato is a tomato to a Pre-Columbian, but it was not invisible to a European, or perhaps gold was more precious to European, but wasn’t nothing to a Pre-Columbian. Tomato was red visually for both, but that color did not have the same connotations on both sides. And a total gesture while eating a tomato was probably -or potentially- different reciprocally, but most probably connected to a complex of many gestures related/attributed to the other, that foreign being.

result may not be the simplest version, but the most perceptually functional one.

2) Timbral layer: Changes in timbral qualities (sometimes as 'timbral gestures') can cause a change on the general impression-effect of the music, also creating emotional and cultural references (especially when considered after changes, retrospectively). Timbral change may be a gesture when it is emphasized (e.g. when it is sudden or extreme in other parameters). In larger settings, when vertical variety and motivic complexity increases, this layer grows into more textural elements. Thus, timbre and texture form a layer which is perceived differently depending on the listening mode. In earlier, traditional terms, this layer can be labelled as harmony and orchestration. Here, timbre is used as a new substitute for harmony, and which is supported by gestural layer.

3) Memory layer: This is the layer where figurative gestures and timbres are processed, filtered, and hierarchically ordered according to their perceptual frequencies, effects, and influences. It is not functionally equivalent to the other two, but it is a ground for and result of their interaction. The memory layer gains intrinsic importance with the repetition of gestures through the musical work or with multiple listenings. The perceptual present rule especially applies for the first listening. Since (long term) memory, enculturation, is the source of expectation; in later listenings, gestures become a part of the line of expectations, creating and reshaping them each time the music is listened to. As a final stage, the conclusion made by the listener depends on the inference from gestural and timbral changes and their combined image through memory shapes the musical form.

The idea of memory layer is consistent with the Schenkerian thinking of form. As an example, Schenkerian analysis of a sonata (regardless of what happens in exposition and development sections, recapitulation) concludes the (most perceivable) structural degrees (of *urlinie* and *bassbrechung,* thus in total, *ursatz*). The difference in proposal is that it depends on the gestural qualities; recapitulation (or any conclusion part) may be less prominent than some earlier sections of the musical work.

In short, a method summarized on the above idea is based on gestures and timbres, processed, and structuralized through memory. It does not take tonal relations as the

basis of the analysis, but the liminality in continuous processes, breaks and strong, unexpected moments and their process. It can be developed and modified to be used for any musical work from any period.

Concepts that will structurally support the proposed approach will be explained in the following sections.

4.2 Timbral Object and Network Analysis

Formal timbre, as described in 2.4.1 is the beginning point for a timbral network analysis. Timbral totality is a signature of any given work, and it should be ranked according to a fixed reference. Measuring average values for some basic timbral qualities (like spectral centroid, spectral flux, etc.) might help to achieve this. However, these values do not mean much without examining their real implications in musical settings and auditory effects. Therefore, reference values defining the complete scope of timbral parameters should be set, in order to get a workable grid.

The levels set in 2.4.1 defines the scope of analysis, varying from a timbral unit to a formal timbre. The descriptors mentioned in 2.4.2 are explained in more detail in the following topics. They will be applicable for all timbral levels unless indicated.

4.2.1 Source identity

An instrument name is a direct timbral descriptor. When talking about orchestration, or in a score analysis, names signify timbres of the instruments, not physical bodies. Therefore, it is important to acknowledge that these names already set a common language for timbre. However, their functionality needs to be considered in certain issues.

Name as a descriptor, by default, refers to solo instrument being played with no modulation, [66] only using the traditional sound producing method. When the timbre is modulated without recognizability is being lost, there would be two reasons: (i) Result of the modulation does not change the spectral components essentially but add another layer of a musical parameter. For example, a vibrato or tremolo on a string instrument changes the timbre slightly, but it is still recognizable due to small amount of change

[66] Modulation here means any kind of gesture that modifies the timbre of the instrument (See 3.2.1).

in spectral components of the sustaining sound. A tremolo would indicate an extreme increase for the emphasis of the attack portion of the timbre, however, sustain portion is always enough to keep the initial qualities. (ii) Some modulated timbres are already considered as a part of macrotimbres of those specific instruments. 'Bowed string' is a timbral type naturally, but 'pizzicato string' is one too.

A name descriptor cannot be an exact equivalent for an instrument group timbre. However, even if it is not possible to create a complete model of a specific instrument's timbre with a group, many compositions use this idea to some extent.[67]In addition, such attempts can also metaphorically position the instrument group as an extension of a certain instrument, or a soloist. A timbral continuity can be sought and notated in such situations.

Last issue about the source identity is the cultural connotations. As mentioned before in 2.4.2, cultural aspects of timbre analysis are inevitable; and as stated in the introduction of this chapter, the principles of the method should include to "avoid semantic analysis and concentrate in physical, physically definable attributes". The main idea behind this approach (which might look like a contradiction) is the assumption that the effects of cultural aspects of any musical structure do not shape the form, but add another layer of meaning that is not directly relevant to the form perception. Aforementioned issue of instrument names and their auditory connotations resulting in timbral types is not related to this issue, since (i) aim and scope is about Western art music tradition and (ii) Western art music and its instrumental collection is the most universal one at the moment.

Conventional instrument abbreviations are applicable for the notation of this descriptor (See Appendix A). Instrumental combinations and ensembles can be derived and modified through this notation.

4.2.2 Registral region and spectral region

The 'pitch' register of a certain instrument is a part of its macrotimbre. Here, pitch should be taken as in Schoenbergian meaning, as a timbral dimension (See 2.3.1). So, pitch, being perhaps the most frequently used parameter for musical sounds, is also an

[67] Gerard Grisey's *Partiels* (1975) is an iconic example for timbral modelling. Composition uses a timbral model derived from a low E2 on trombone.

indicator of timbral qualities. The registral region used in a certain section or complete musical work is like a mark point that the spectral qualities of the formal timbre refer to.

The sum of the complete instrumental collection ranges sets the total register of activity. Musical registers are already labelled in tradition and firm descriptor for analysis purposes (See Appendix B).

Any musical work that uses a certain part of this register can be classified accordingly. This 'class', which can have different implications for different instrument combinations, then should be investigated in its relation to spectral region.

Spectral region can be seen as the equivalent of registral region in spectral content. It is set by the border frequencies of the sound spectrum. Its inner distribution and balance are the topics of spectral envelope and microtimbral details. However, spectral region, especially when combined with registral region, can point to some basic timbral features of formal sections. The most basic relation between registral and spectral regions is the applied energy resulting in dynamics, loudness, thus, the change in other spectral parameters, such as centroid and flux. Since spectral features of any kind of music cannot be easily detected from a score, an indicator for spectral region can be the average dynamic level of the instruments used. As stated in 2.3.1, any timbral region or object will get 'brighter' when more dynamics are applied. Depending on this factor, spectral region of a work and its inner sections can easily be labelled without having too much emphasis on quantitative spectral properties of each instrument. Spectral region is also directly affected by the modulations applied on instruments and playing techniques. However, their representation as separate inputs may be needed, due to their varying effects across instrument families.

Lastly, spectral region can also be referred in accordance with its proportional relation to hearing range[68] when labeling the formal timbre of a work. This and alike analyses have to be computer and/or recording based.

[68] Roughly 20 to 20k hertz frequency interval.

4.2.3 Envelope

Analogy between a timbral unit and an overall timbre is not metaphorical. In fact, any timbral unit stretched over time, or any compound timbre compressed into shorter duration serves the same effect perceptually. However, the resolution[69] of possible analyses differs for each timbral object. Because of this, it is necessary to follow certain top to bottom steps to reach consistent results.

At this point, I propose a practical interlevel analogy of timbral unit and larger timbral objects:

- Timbral objects (including timbral units) can be examined taking the attack-decay-sustain-release (ADSR) envelope as a model.
- Timbral classes (including harmonicity, richness, fullness, spectral distribution) and their combinational results, timbral types (pitched/unpitched, beating/beatless, pure/noisy, bright/dark, etc.), can be used for each of stage of the ADSR envelope. It is also possible to device an MDS-like system for the needs of specific, work or style-based analyses.
- Any stage of the envelope can be absent in a timbral object (physically or perceptually). Nonetheless, physical absence of some stages will be 'restored' by the cognitive mechanism; e.g., absence/presence of attack portion does poorly affect the recognition of the timbral source (Donnadieu, 2007, p. 294) and listeners mostly categorize the attack stage in abruptness/graduality. As stated in Hall & Beauchamp (2009), "distinct performance differences along a rise time dimension are possible in the absence of a fixed category boundary between instruments with abrupt and gradual onsets." (p. 18)

Loudness and duration of each stage is essential to the timbral composition. They should be notated in a simple way, preferably with a vector-like way (e.g., for a timbral unit: ⟨0.2″, 0.1″, 4″, 2″⟩; and for a bigger timbral object ⟨5″, 0, 1′20′, 10″⟩). On a score-based notation of envelopes, assuming a tempo is already given, note lengths can be used to create similar vectors (e.g., q=80, unit a ⟨4, 0, 12, 3⟩).

[69] Resolution meaning the number of layers, levels, and details.

4.2.4 Object classes and types

The approach presented in the previous section requires a simple, functional set of relations between timbral classes resulting in timbral types. Each class and their resultant types will be explained briefly.

- Harmonicity: The level of harmonicity (partials being whole multiples of the F0) results mainly in timbre being pitched/unpitched types. Pitched objects, from a sine wave to instruments sounds and polyphonic combinations, and complex chords,[70] noise-pitch combinations, which can either turn into another timbral object with an F0 or to new noises, depending on the input.

However, a reconsideration is necessary when using concepts of harmonicity-inharmonicity in a quantitative context. As already stated in 2.4.1, harmonicity should not be put in a principal functional position. Recent decades have witnessed the increasing inharmonic timbral objects in contemporary art music and the 'effect' has already passed the limit where it created a binary axis of perception. A simple and important statement on this is made by Julian Anderson in 'Spectralisms 2019 International Conference'[71], saying "I don't really believe in the notion of inharmonic spectrum" and citing Horațiu Rădulescu saying "They don't exist. […] If they are there, basically there's some low fundamental you don't hear" (Anderson, 2019, 24:01). Therefore, it would be an option to use this concept as an indicator for different/similar levels of focus, or possible colors interpretations of a spectrum; rather than a rigid, solely quantitative parameter.

- Richness: The total number of partials in a certain spectrum makes it rich or sparse.[72]
- Fullness:[73] A certain selection of the harmonics present in the spectrum defines its fullness. Full harmonics mean full sound, a portion of them results in hollowness (e.g., clarinet's spectrum having -or dominated by- odd harmonics results in a hollow sound) Thus, this actually becomes about the deviation in

[70] As discussed in chapter titled 'Chords with six or more tones' in *Harmonielehre* (Schoenberg, 1978, p. 411).
[71] This second edition of the conference took place on 12-14 June 2019 in Paris during IRCAM ManiFeste festival.
[72] The terms rich/sparse are taken from Cogan (1984), in which he chose binary descriptors for timbre analysis. He also labels them as negative forms (low energy) and positive forms (high energy).
[73] The terms full/hollow are taken from Lavengood (2017).

harmonic numbers. Any sound may have a large number of harmonics but the balance of them can make it sound hollow.

- Spectral distribution: The distribution of the spectral content over the totality of the spectral region results in bright/dark types of timbres; bright timbre being balanced towards higher frequencies and dark being the opposite, i.e., position of the spectral centroid in total spectral region. This means, brightness of the spectrum needs to be decided by using the spectral balance (or shape) between F0 and the highest partials. This is one of the crucial principles for defining the overall and formal timbre. In addition, as mentioned in 2.3.3, brightness also increases in parallel with loudness. This class can easily be categorized and notated by using (L)ow-(M)id-(H)igh-(E)xtreme notation, resulting in types of L_E, L, M_L, M, M_H, H_E, L_E-H_E, L-H, M_H-L, etc.

Two other type categories can be added to this list including compact/diffuse and beatless/beating.[74] In this proposal they are used in following definitions:

- Compact/diffuse refers to the width of the spectral bands around partials. Natural sounds fall into diffuse type and electronically created sounds are more compact. This can be seen as another explanation, or equivalent of spectral flux.
- Beatless/Beating is sometimes used as 'pulsating', and it indicates the presence of pulses that are caused by strong harmonic centers that are in close proximity. This closeness can be explained by aforementioned critical band for the given frequencies; if the difference between two main frequencies are in a perceivable (i.e., 'countable' in this case) they create a beating. This type can be seen as a magnified, extended version of the previous, aesthetically.

Unpitched types (i.e., noises) can be categorized using the same logic and using common colored noise types (Noise types in continuum are: Pink noise, brown noise, white noise (and grey noise), blue noise, and violet noise). Normally noise types do not need spectral region notations, however, sometimes in some in-between situations, as mentioned above, (i.e., when there is an audible fundamental in the noise mass) they can be used in combinations. It is important to remember that in instrumental settings, these noise types do not mean much, because instruments are not capable of

[74] The terms compact/diffuse and beatless/beating are taken from Cogan (1984).

creating and filtering out any frequencies with such precision. So, if they are used for acoustic instruments, they will be simply metaphors, or approximations of them. However, it is possible to use them in direct definition in hybrid (electronic sounds-acoustic instruments) compositions.

These classes and types are applicable to all stages of ADSR, but they may result in slight differences in their effects due to durations and amplitudes. For example, in a timbral unit, a very short attack time renders the harmonicity of the attack portion meaningless, making it percussive, since under 200 ms it is not possible to hear timbral qualities. Nonetheless, similar conditions get easier to analyze in larger timbral objects.

Last type descriptor, modulation, should be based on a list of techniques that are common and easily definable. If there is a technique of modulation which is very specific, it should be added as a 'specificity'.[75]Some modulations can be traditional instrumental techniques like changing positions, pressure and movements (tasto, ponticello, col legno, overblowing, vibrato, etc., see 3.2.1); or can be applied by outer forces like amplification, electronic audio equipment, extra resonators and modulators; and more importantly, the space.

4.3 Gestural Network Analysis

In traditional ways of form analysis, harmonic structures are taken as the basic units and general form is evaluated through connections of these. Even though this approach is valid for the material, compositional aspect of the musical work, it is not enough to explain the perceptual effects on the listener's side. Gestures come into play at this point. Studying musical form is a study of prolonged gestures and their conclusions, where they are detached from the material world and exist only in the musical dimension; and gestures create continuity in music (Scruton, 1999, p. 341; Scruton, 2009, p. 54; Hatten, 2004, p. 240). Even if we refer to conventional meanings of music and form, this idea of 'prolonged gestures' can still be valid for the holistic definition of music as in this proposal.

[75] The term appears in McAdams et al, 1995, and in later articles by McAdams.

4.3.1 Gestures and timbral gestures

There are two ways to interpret the notion of gesture in this context: (i) as a network that is integral to timbral coherence, where the gesture is directly a part of timbral objects; and (ii) as structures that 'do not have timbral properties'[76], nevertheless, interfere with the perception of timbre and its formal network.

Gestures themselves may not have timbral significance but may create formal sections. Timbral gestures can form formal sections, but with repetition on gradual changes (depending on the duration) may have the opposite effect. For example. transformation of a timbral unit or object can be considered as a timbral gesture in the middle ground or foreground analysis, however, this type of timbral gesture has the potential to create sections not with separating structures but combining them into a bigger section.

Two other important notions for gestural bodies is 'transposition' and 'mutation'. Transposition inevitably has the association of changing the reference point of a pitch sequence while keeping the inner intervallic relations; and this definition still applies here. However, rhythmic and durational features of the gesture do not always have to be kept exactly. This is especially valid when a gestural body does not show any significance in terms of its rhythmic pattern, which is highly possible when the meter of the music is not emphasized in a traditional manner. Thus, transposed gesture can be augmented, diminished or 'mutated' if the salient characteristics in other aspects are kept.

Mutation -as it is used in this thesis- indicates slight changes in several aspects of the gesture. Changes in similar pitch intervals, inversions and retrogrades[77], addition and subtraction of certain elements can be given as examples. Yet, again, these changes must not be fundamental, must not appear as new gestures. A mutation should always be recognizable auditorily. Several examples of mutation will be presented in Chapter 5.

---

[76] Every musical structure has timbral properties, here it is intended to mean that for parameters other than timbre are more foregrounded than its timbre. This happens due to context it exists in.

[77] For inversions and retrogrades, in small scale gestures only, if the parametric relations are easily recognizable.

4.3.2 Temporal Considerations

By definition, a gestural network needs to take shape in the second level, since all prolongations and transformative relations of unit gestures are already categorized in temporal dimension. Nonetheless, analysis and evaluation of gestural elements has to be done in both directions in this hierarchy. The structural hierarchy of different level gestures may not always be kept in their interaction. For example, a gesture having a motivic function needs time. So, motivic development of any gesture is vulnerable to any interrupting first-level gesture that is significant enough. In addition, the strong gestures that can act as timbrally unifying factors (See 2.2), would also be a distracting factor that breaks the horizontal continuity. This way, horizontal continuity of one line hides the continuity of another horizontal evolution of a timbre. Cone's idea of 'music as gesture' that is presented in 3.2.3 can be followed and developed to serve this line of analysis.

Unit gestures can also create networks over the duration of the work without transforming into motivic material; however, these networks may stay as 'commentaries' on the real subject (motif). A unit gesture needs to be focused on, be presented in a suitable texture, to reach a second-level network (See 5.1, 'two-part gestural body').

4.4 Significance and Markedness

There can be type-related similarities between timbral units and higher level timbral objects, even a formal timbre (as in possible similarities between gestural levels). Yet, their individual properties are not enough to categorize and formalize them in a structural form. They need to have gestural significance or a certain level of markedness. The term 'markedness' refers to a condition of contextual asymmetry that any gestural feature of a timbral object obtains. Lidov explains the origin and the meaning of the term in the foreword of Hatten (1994) as:

> Markedness theory proposes that wherever humans draw distinctions (right/left, man/woman, etc.) these tend to be asymmetrical: One side tends to be more richly evaluated (positively or negatively) and more special; the other, to lend itself more to abstraction and sometimes to represent the divided whole (as was until recently the case with "man" in English). (Hatten, 1994, p. x)

The example given by Lidov derives from the linguistic field and it is a direct generalization of psychological tendency of humans. Considering the functional analogies drawn between gesture and language in 3.1, it is necessary to point out that same logic of marked words would be applicable to gesture, in this case to musical gesture. Therefore, an analysis of markedness in combined gesture-timbre network should take necessary presumptions into account. Inevitably the question occurs: What presumptions?

First topic to discuss the possible answer or answers to this question is the culture and long-term accumulation of signifiers. Hatten (1994) uses major and minor modes in a comparison where he positions minor as the marked one with a narrower meaning (i.e., tragic) and major as the all opposite, comic, with all other meanings (i.e., heroic, pastoral, etc.) (p. 36) His claim is completely valid, considering the book he states this idea is about Beethoven's music. Thus, his analysis, and therefore, his presumptions are genre specific. So, as also stated in 2.3.5, anything that can be considered 'less usual', 'less expected' or 'un-expected' is more significant (e.g. if a triplet is used in a rhythmic context where everything else is duplets, it is marked; a soloist instrument is marked, an instrument reaching for an unusual timbre is marked).

Considering the actualities of contemporary art music, this fixed set of dichotomies would not work; and the plurality of everything in contemporary art music denies the term of genre, in a sense. Therefore, genre-based presumptions should be substituted with work (track, piece, movement, section) based examinations. Even in musics labelled with genres, it might not work,[78] because once the listener listens to the work, context changes. Secondly, re-listening of any work (which is a default musical activity since the wide availability of recording technology) would take the perception of marked gestures and timbres to a more objective state. This objective state is the physically framed, culturally unburdened set of analytical mechanisms. Needless to

[78] Taking the issue of genre to a different field, an example can be given from'black metal'. Being a subgenre of metal, black metal bands use harsh vocals, with their guitars in full distortion, playing quite dissonant or folkish contrapuntal music most of the time. However, there are many examples from the same bands that use clean guitars and vocals, tonal chord progressions, and popular song forms. Thus, many tracks become marked in the context of a 'music album', in this specific genre. The form and analysis of the track becomes independent, because even without the 'expected features' from a 'black metal' track, it is still considered black metal because of the band's identity. Similar cases of non-sound-based genre descriptions can be found even with 'midi compositions', which were still considered 'black metal'.

say, any analytical approach from any cultural perspective can work on the same thing, however, that would serve in the analysis of the culture itself, and not the analysis of music.

4.5 Contextual Considerations

Regardless of any analyses on particular timbre and gesture complexes, one of the important topics to discuss is the contexts that are created by certain conditions that belong to the object of analysis. These conditions can be about many different timbral or gestural level features of the given work; it may be impossible to make a full list of them. These can be seen analogues as timbral 'specificities' that are presented in 2.3.3. Nevertheless, these issues can be approached with a consistent logic that is already presented in the previous sections.

As an example of timbral context, the topic of homogeneity can be given. Homogeneity of instrumental combinations changes the impact of any given parameter; e.g., if we are talking about a string quartet, or orchestra (in old terminology, a whole consort) the common parameters, (e.g. spectral flux, attack times, everything related to spectral components etc.) become irrelevant, since they are almost the same, or continuous in a very gradual way. Thus, this timbral parameter becomes obsolete in timbral analysis of form. However, it is still relevant in overall qualities of timbre for the given musical work. In this case, 'homogeneity' becomes a parameter to indicate how to quantify or qualify the timbre.

Sometimes any object can lose its timbral properties due to context. This could happen when there are extreme differences in some parameters. For example, an object with a short attack and short sustain values in a timbral environment with longer sustain values (basically, continuities in temporal or vertical axes), would be perceived not as another timbral object but more as a texture changing gesture. Nevertheless, a repetition of the same object can bring out its timbral properties too. This also can be considered as an issue of 'markedness', since the dominating characteristics in a certain sonic environment creates its regular 'unmarked' parameters and unusual 'marked' ones.

The same idea is applicable for gestural complexes. Aforementioned issue of gestures being functionalized as motifs or figures, can change the context of analysis. For

example, when a musical work has little amount of conventional gestures, a gesture that would be considered as insignificant can be essentially important; or the opposite, like a music with gesture-rich complex texture may need a 'anti-gesture' kind of non-motion or silence as a structural element.

Contextuality is rather a checkpoint that consists of several aspects to consider, than a fixed list of points. It has to be renewed for each work, depending on its timbral elements.

4.6 Re-Evaluation of Form in a Fictive Case

This analytical approach can be summarized in two basic steps:

Timbral features of the total**:** First category is about the overall characteristics of the timbre, or, characteristics of the 'formal timbre', spectral realm and other overall qualities. In popular terms, it is the equivalent of 'the sound', the coherence that forms the totality work, shapes the outer boundaries of the timbre. For example, (if we are talking about a recording) a track that is recorded, mixed and mastered in a specific way (i.e., in a specific set of parameters), makes one to recognize the overall effect in identification of that work. In a case of recording, one would know a certain song from a certain album at an instant, or, even if that listener does not know the work, artists, album, period etc. may be guessed. In the case of an acoustic experience of a 'contemporary classical' ensemble work, it is more difficult to identify it as belonging to a certain style, school etc.; but timbral coherence at this level would not work easily depending on this style and general characteristics. However, in this case, it is still possible to analyze and categorize formal timbre by analyzing the used spectral totality comparing to the auditory boundaries by creating a spectrograph showing the total weight distribution of frequencies of the whole work. As a subcategory timbral homogeneity of the ensemble becomes a parameter to how to approach timbral sections, transformations etc.

Timbral features of the sections that shape the inner divisions/contrasts (which we call as 'the form')**:** Second category is the level of continuum and segregation in the temporal axis throughout a given work. Which means, timbral coherence is sought in vertical axis to categorize the timbral regions and in temporal axis to see the similarities, differences, repetitions, abruptness, or gradualness.

This step must be completed by taking many factors into account. Taking the first step as a reference, a restricting frame (the spectral realm that shapes the totality), each section is determined, from bigger to smaller.

- Firstly, given sections, like movements, specific sections that are already indicated in the score.
- Secondly, an overall analysis of gestures and gestural development. The breaks and changes through different levels of gestures. Again, starting from bigger to smaller. In addition, depending on how often a gesture occurs, repeats, or gets emphasized. Same logic is applied to all 'timbral gestures' on their frequency of occurrences (See 4.3.1)
- Thirdly, classification of the timbral qualities of the sections (that are the resultants of the previous steps):
    - Firstly, multidimensional vector types that are determined for the sectional characteristics are calculated. Result is a combination of vectors that should be sought in all sections.
    - The results of the first timbral analysis, the comparison of these vectors, give us a rough sketch of the overall form.
    - However, re-consideration of the form is needed, in order to see if there are timbral features of these sections that reach beyond and create new relations with other sections, possibly changing how we see the form in general. This can be seen as realizing the timbral parentheses.

A sample scheme depicting the relations and the mechanism for analysis of a fictive musical work will be presented (Figure 4.1). Here, some fictive musical gestures and their relational network in a fictive musical work are reflected onto the scheme, by using varying sizes of vertical lines to refer to their level of significance. This gestural layer is reflected to memory layer together with timbral layer, where their larger structural relations and frequencies are compared and evaluated in order to discuss their combined effect.

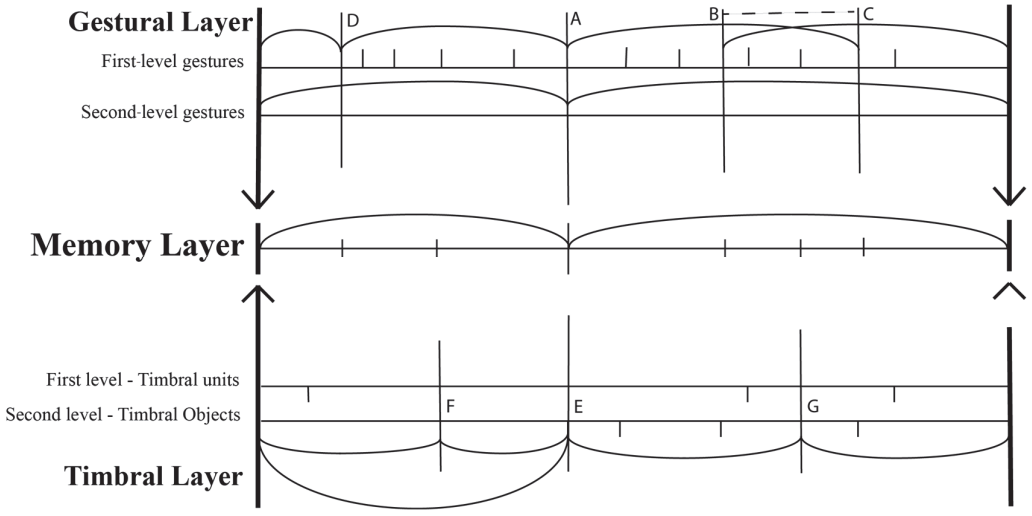

Figure 4.1 : A sample scheme depicting the three layers for analysis.

In the above figure, gestural hierarchies are presented in two levels, first and second. In addition to this, every level has two sub-levels[79] of significance that are indicated by shorter and longer vertical lines reaching upwards. In the first-level gestural layer, point A represents a moment where first and second layers coincide and make a strong affect that is conveyed to the memory layer. [80] The region between B and C represents an overlap of two gestural regions. Beginning point of the region will be significant since a new gesture is introduced. This new gesture would potentially change the perception of the already existing one, thus creating a formal point. However, end point of the region would also be important since disappearance of the first gesture has the potential to change the perception of the second gesture. Point D represents a strong first-level gesture that cannot reach a greater formal importance because of the strong continuity in the second level; however, it still represents a lower level of significance. Shorter lines following the first one can be thought as the mutations of the first gesture, or individual ones, yet, coherence in the second level is stronger than the first. Thus, all four points are reflected onto the memory layer, however in this case, A is the most significant one to create a formal juncture because it coincides with another point coming from the timbral layer.

Same structural approach is applied to a fictive timbral layer in the lower part of the

[79] Any musical work can have more levels, depending on the gestural complexity and musical context.
[80] E.g., a moment when a strong unit gesture coincides by a climax point for several aspects, or used to make an ending stronger.

scheme. A combination of first-level and second-level timbral events are listed in two layers. In this case, the points in first level are only reflected to the memory layer only if they create a second-level effect or coincide with one. Thus, only point E is the most significant one because it coincides with point A on the gestural layer. Points F and G are also reflected onto the memory layer, but they have less significance compared to E.

In this scheme, memory layer represents the accumulation of all moments created through significances that are rooted in both short-term and long-term. As stated by Pressnitzer and McAdams (2000, p. 55), if a new stimulus is similar enough to any previously categorized representations, it will be considered as a member of the same category. They refer to the terms of 'perceptual present' and 'working memory' with a similar approach and give references to the aspect of cultural familiarity (Pressnitzer and McAdams, 2000, p. 54-56). Schenker (1977, p. 128) points out a comparison by contrasting the memory of the background (here presented as the totality of the form) and improvisation, which basically brings up the aspect of familiarity and expectancy, although it represents a lesser degree.

It is also necessary to add that, even though familiarity in a certain style can sometimes bring advantages in perception by directing the focus to certain parameters in music; expectation-violating concepts can have more salience and advantage to be propagated. As stated in a categorization by Snyder (2001), musical forces can attempt "to exploit long-term memory by building up hierarchical and associative mental representations of large time structures" or (as demonstrated by empirical data in Parubanova, Shaw, McKay and Xygalatas, 2014) "sabotage recognition and expectation by frustrating recollection and anticipation, thereby intensifying the local order of the present." (p. 234) Finally, when the features of musical events and their combined effects are examined by using a set of parameters specific to a certain genre or work, memory layer is shaped.

5. SOME PROVISIONAL APPLICATIONS

Developing a consistent method using the information gathered and ideas presented in the previous chapters needs another extended study. In this chapter, following analyses will be provisional ones; they are initial considerations of the principles that are explained in the previous chapters. Various methods and their applications can be valid from a similar perspective. However, these will be presented as author's initial attempt to develop an applicational approach. It is also crucial to remember that these analyses also depend on the basic knowledge of instrumentation, macrotimbral properties of instruments and their dynamic capabilities.

Two works will be partly and briefly pre-analyzed: Anton Webern's "*Fünf Stücke für Orchester*, Op. 10, No.1" as an example of early post-tonal period, putting an emphasis on timbral listening; and "Threnody I: The Night of the Electric Insects" from George Crumb's "Black Angels" for electric string quartet, which is known for its use of extended techniques, contrasting timbres and extreme registers.

5.1 Anton Webern, Op. 10, *Fünf Stücke für Orchester* (1913)

Anton Webern's (1883-1945) *Fünf Stücke für Orchester*[81] (Five Pieces for Orchestra) arrived a few years after Schoenberg's well-known Op. 16 with the same title. Utilizing the technique of *klangfarbenmelodie,* Webern creates his own version in a brief work, lasting around six minutes in total. Two of the pieces were composed in 1911 and the other ones in 1913 (Angerer, n.d) and he himself conducted the premiere during ISCM Festival of 1926 in Zurich (Bailey, 1998, p. 121). Webern originally gave names for each part: *Urbild* (prototype), *Verwandlung* (transformation), *Rückkehr* (return), *Erinnerung* (memory/remembrance), *Seele* (soul), respectively (Johnson, 2006, p. 123). However, first printed edition of 1923 dropped these names.

Work is notated for a small chamber orchestra, all solo instruments:

[81] For the full score of the piece, see Appendix D.

Flute (doubling piccolo), oboe, clarinet in B♭ (doubling bass clarinet in in B♭), clarinet in in E♭, horn in F, trumpet in in B♭, trombone, harmonium, celesta, mandolin, guitar, harp, percussion (glockenspiel, xylophone, cowbells, low-bell, triangle, cymbals, side drum, bass drum), violin, viola, violoncello and double bass.

Instruments are used in a very frugal way, most of the pieces lacking a few of them (e.g., no's 3 and 4 do not have either flute or piccolo; bass clarinet is only used for no. 5, shortly). String and brass instruments are mostly used with mutes, consistent with the low dynamics character of the work.

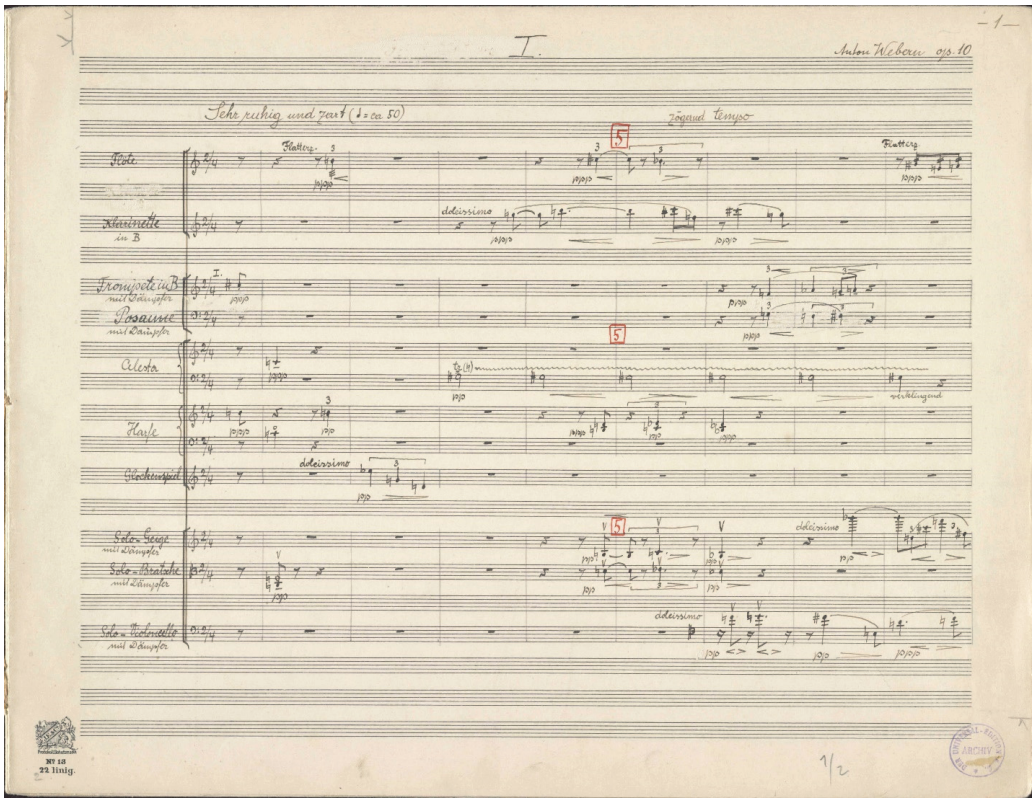

Figure 5.1 : First page of the full score, Webern Op. 10 (1913) (autograph manuscript).

In a program note written for *Festival d'Automne à Paris* (1980), Pierre Boulez described Op. 10 of Webern as a work where "tonality no longer exists; it is no longer

the one to organize the language and structure" [82] and he emphasizes the lack of repetition, even the "repetitions of timbre". Despite the lack of complete gestural and timbral repetitions, moments created through timbral contrasts and continuities exist throughout the work. All five pieces are quite short, but they are also available to be examined in smaller formal sections.

A general discussion and partial analysis of first piece is as follows.

Sehr ruhig und zart (very calm and soft, q=ca. 50), starts with a two-part gestural body, where the parts that are separated both by timbre and register but fall into one gestural unity due to total duration and with the help of dynamic crescendo in the flute. This gestural body directly presents itself with a potential of creating second-level connections due to present texture, where it is presented as sole musical material. The issue of duration is directly related to the definition of 'expressive unit gesture' that is explained in 3.2.2. This short duration around 3 seconds for each part, provides the pitch elements a gestural unity. Timbral aspects of the given instruments stay relatively stable, keeping it timbrally coherent. It is applied by keeping trumpet and viola muted; and other instruments except flute are low attack-time instruments. To support this observation, the changes in spectral centroid and spectral flux for the whole piece are presented in Figure 5.2 (in colors red and green respectively). Rectangular area shows the area of first two gestures that take place between 3 s – 9 s in the used recording. [83]

The initial, pitch related gesture presented in the first bar is active throughout the piece, sometimes modified on minor and major second intervals and rhythmic augmentations. The summary of these gestural transformations with octave equivalence applied can be seen in Figure 5.3.

---

[82] Original quote in French: «Dans les pièces de l'opus 10, la tonalité n'existe plus ; ce n'est plus elle qui organisera le langage et la structure de l'œuvre ; face aux difficultés nouvellement surgies, Webern s'en tient aux formes extrêmement brèves, où tout est essentiel, d'où est exclue toute répétition, même ce que l'on pourrait appeler les répétitions de timbre.»

[83] The software used for analysis is 'Sonic Visualiser'. It is a free, open-source application for developed in the Centre for Digital Music at Queen Mary, University of London.

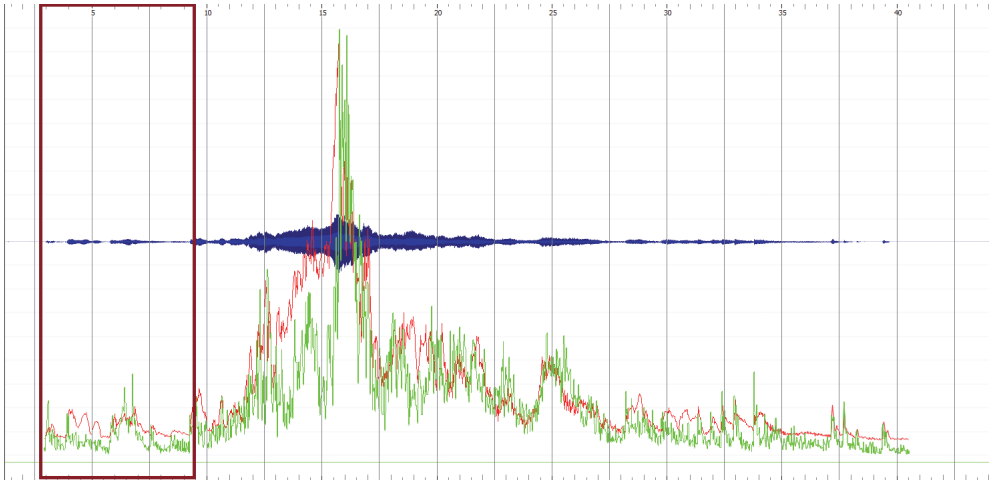

Figure 5.2 : Proportional graphic representation of the overall spectral centroid and spectral flux for Webern's Op. 10, No. 1.

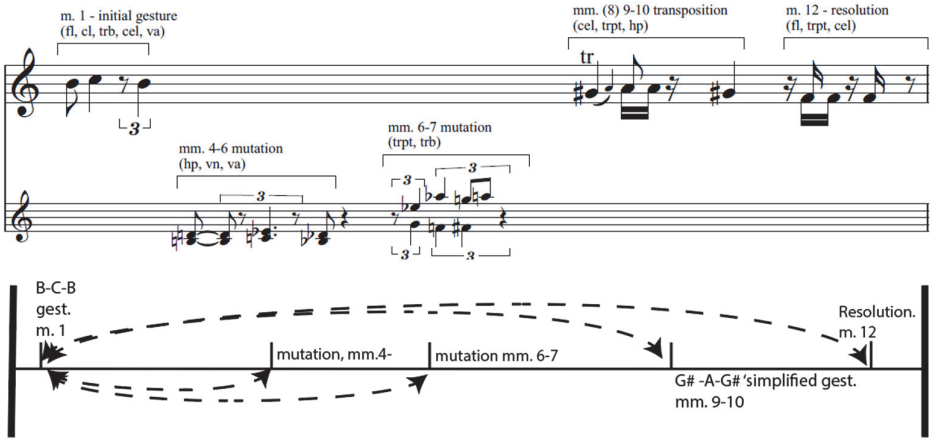

Figure 5.3 : Transformation of initial gesture and its summary on the gestural layer.

A possible discussion appears at this point about whether to see above gestures in the first level or in the second level. Depending on the definitions that are presented in Chapter 3, gestures in the above staff belong to the second level since they are combinations of pitches performed with several separate instruments. Consistent with the idea of *klangfarbenmelodie*, they exist both in vertical and horizontal axes and these separate instruments can poorly serve anything as first-level gestures. Therefore, they will be accepted as second-level gestures in the following schemes. However, it is also an option to consider them as first-level gestures if it is accepted that they present a unified, melodic unit, which can clearly be perceived as a unit gesture. Additionally, following this idea, it can also be noted that initial gesture and its transposition and resolution follow a descending line of minor third intervals (B, G♯,

F, thus, a tritone) which also can be seen as an arpeggio, creating another second layer relation.

Mutations of the initial gesture between mm. 4-7 already fall in the definition of second-level gestures since they present a complex of related but nonidentical gestures.

Second part of the gestural body is the descending line played by glockenspiel. It is reminded later in mm. 10-11 by harp, and it also serves a similar function at that point where timbral features are important. It is also significant that two gestures share the same interval relations in totality. They use different registers but use ninth intervals (E♭- D, E-D) in maximum range, similarly. One other significant (melodic) gesture appears in violin line at m. 8. This gesture is made significant again through its register and slight dynamic ascend. Following bar also has glockenspiel for the second time with a one note gesture, however, this gesture is significant because of its timbral qualities and will be mentioned below. All of these gestures, by definition, fall into the first-level gestures. These additional gestures and their replacement in the gestural layers can be seen in Figure 5.4 and Figure 5.5.

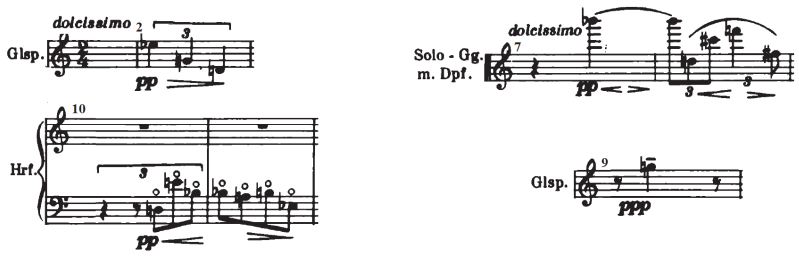

Figure 5.4 : Additional gestures, mm. 2-10.

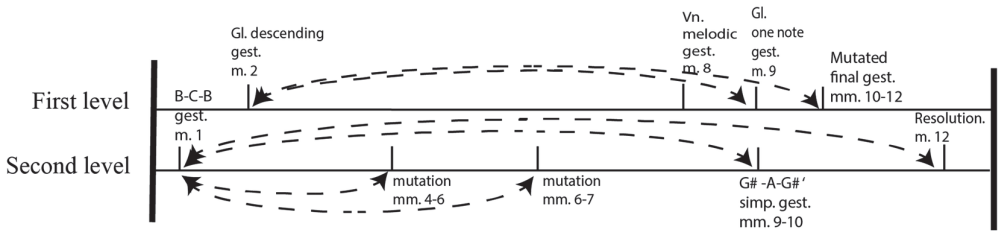

Figure 5.5 : Additional gestures added to the gestural scheme.

Despite the relatively frequent appearance of initial gesture, it does not perceptually serve as the same at every moment it appears, due the octave changes and harmonic combinations. The instrumental summary and pitch relations of the first gestures can be seen in Figure 5.6.

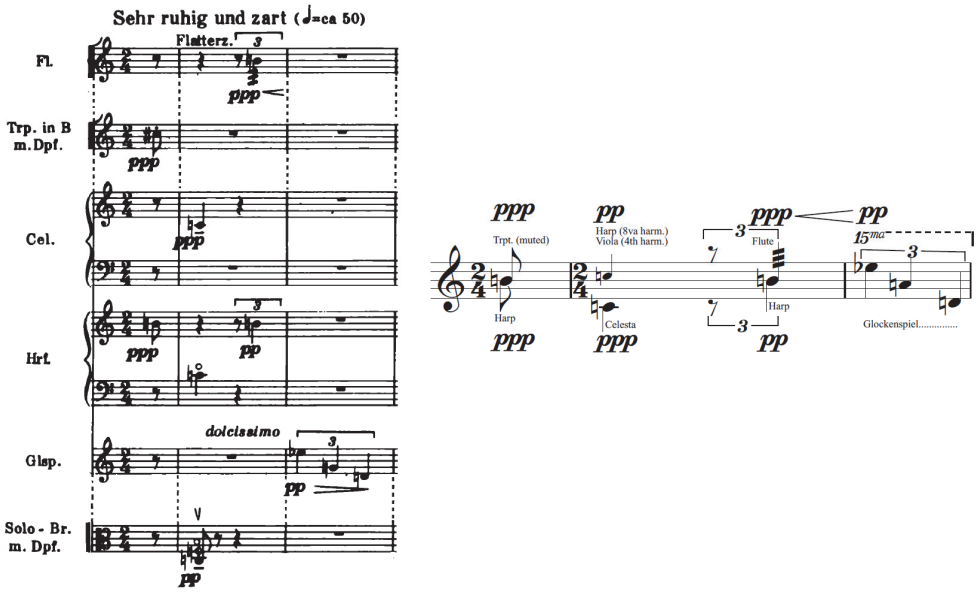

Figure 5.6 : Mm. 1-2 (adapted original) and its reduction.

A way to analyze the effect of glockenspiel's entry would be a 'timbral interruption' that leads to a developing timbral object. This idea can be supported by (i) the registral difference glockenspiel has, which makes it clearly distinguishable and (ii) its timbral connection between glockenspiels descend of E♭7 - G6 – D6 and celesta's G♯3 trill figure that is sustained for the next six measures. Because of the registral difference, the suspending notes from descend of the 'non-dampened' glockenspiel function as partials of celesta's G♯3 trill. G♯3 trill is between G♯ and A and because of this minor second interval quality and sustain, it is a 'pulsating' timbre type. The spectral flux created by this pulsation, serves as a basis for the developing timbral object of the next seven measures.

The development of the timbral object of mm. 3-9 can be reduced to several regions with basic macrotimbral features and their registral continuum.

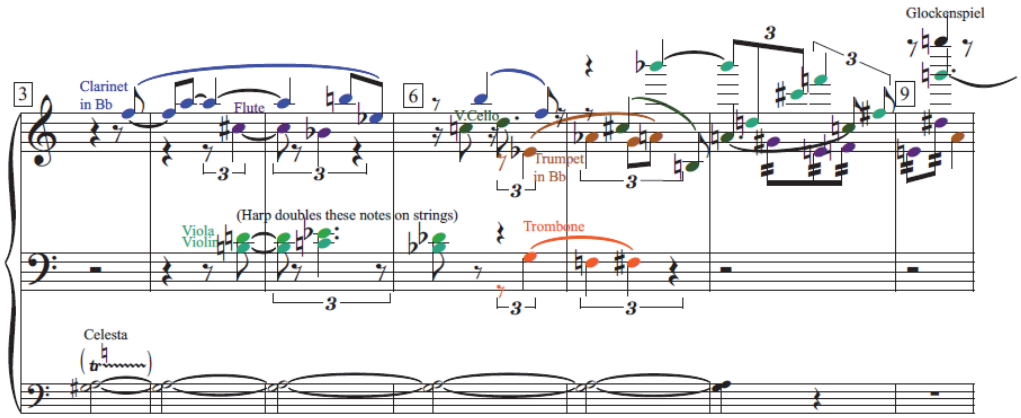

Figure 5.7 : Reduction of mm. 3-9.

As seen on Figure 5.7 (colored in similar tones for instrument families), registral region for the totality gradually evolve during these seven bars, until the glockenspiel hit the high G and signaling the end of the object. The effect of instruments that are grouped in families acting in parallel registers, makes this transformation easier. The clarinet line covers the upper registers and contributes into the growing timbral object, briefly shifting the focus from celesta to itself. This shift is also clear with the ascend and the descend of the spectral centroid in Figure 5.2. However, this clarinet line is a relatively simple one, perhaps, which should not be considered as its gestural features when compared with the general characteristics of the moment. Only significant registral jump happens in violin line in mm. 7-8; and this change also marks a point where violin serves a gestural purpose as it was stated in Figure 5.4 and Figure 5.5. A simplification of the registral transformation can be seen in Figure 5.8.

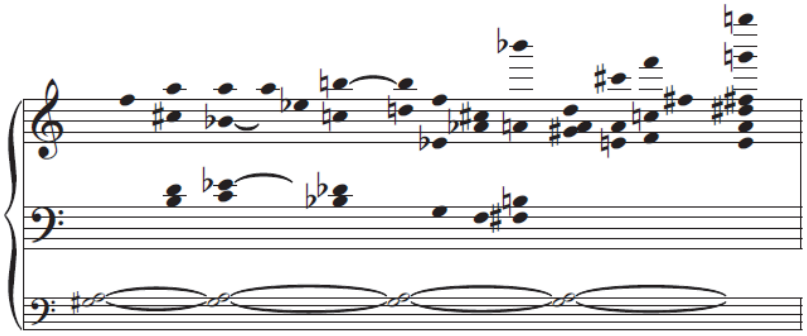

Figure 5.8 : Simplified registral transformation of mm. 3-9.

One of the most remarkable features of the first piece is that its loudest individual dynamic is ***p*** and softest is ***ppp***. Thus, the overall timbre is not significantly modified

by dynamics during these forty seconds. This also helps to see that the total dynamic level is built with the number of instruments playing at the same time and not any sudden gestures.

Glockenspiel's high G7 on bar 9 (a) signals another interruption that ends the evolving timbral object, yet it does not start another one. Violin's accompanying G6 (b) serves as an extension of glockenspiel, rather than an echo of the violin gesture on mm. 7-8 (c). Here violin's gesture is chosen especially, due to its rapid features when compared to other melodic lines that do not show any significance. Thus, violin melody here is 'marked'. Glockenspiel's simple gesture is followed by a transposed, orchestrally simplified partly augmented version of the C-B-C gesture in the beginning, as G♯-A-G♯ (d). This moment prepares the end of the piece and harp's harmonics (e) version of the 'interrupting' initial glockenspiel gesture (m. 2) which is already signaled by one note glockenspiel gesture (a). Last gesture appears as a cadence of F's (f), with the inclusion of harp's gesture, as another mutated final gesture, which is the replacement for the diminished one in mm. 9-10.

The totality of relations that are mentioned above can be seen in the full score presented as Figure 5.9.

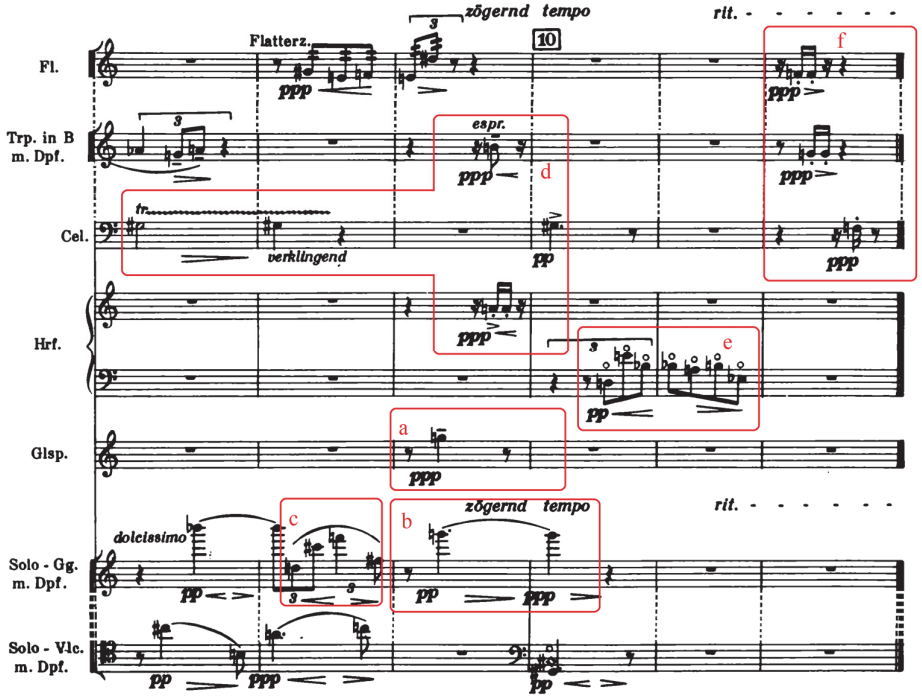

Figure 5.9 : Adapted score of mm. 7-12.

Concluding the analysis of the gestural and timbral relations in the piece, the notation method that is proposed in 4.2 can be presented as follows. Only the main, formal regions are given with their most significant gestural units/objects.

q=50, (fl, cl, trpt, trb, cel, hp, gl, vn, va, vc), mm. 0-12
reg. "a": gest. "i", (trpt-m, hp, cel, va-m, fl), mm. 0-1;
reg. "b": gest. "ii", (gl), m. 2;
reg. "c": t-o "x", (cel, cl, fl, hp, vn-m, va-m, vc-m, trpt, trb-m), mm. 3-10, (2, 0, 3, 2), M_L→M_L-M-H→M-H→M-H_E;
gest. "iii": (vn), m. 8;
gest. "iv": (gl), m. 9;
gest. "i-2": (cel, trpt-m, hp), mm. 3-10;[84]
reg. "d": gest. "ii-2", (hp); mm. 10-11;
reg. "e": gest. "i-3", (fl, trpt-m, cel), m. 12.

In this notation, first line indicates the tempo, instrumental setting, and total number of measures. Measure number '0' is used to indicate the anacrusis, otherwise it would start with '1'. Following lines indicate formal regions (abbreviated as 'reg.') where they are either labelled with gestures (abbreviated as 'gest.') or timbral objects (abbreviated as 't-o').[85] Certain gestures that exist in the region labelled as a timbral object (reg. "c") are notated as they are contained in the region by an indent.

These regions are defined by their instrumental combinations that are ordered by their entries from first to last and on the score from above to below. If there is any modification to the instruments, they are indicated (e.g., some of them are muted and they are notated with an '-m' suffix). Re-entries of instruments are not notated unless they create another object or region.

Timbral object "x" also has two more components: envelope and spectral distribution. Envelope follows the approximate dynamic levels and their continuity. Here, it is notated in number of measures, however, it is possible to notate it in real durations, depending on a recording. Spectral region change is indicated with [L]ow-[M]id-[H]igh-[E]xtreme notation given in 4.2.5. With the addition of second-level gestural regions and the whole timbral layers, Op. 10 No. 1 can be summarized in Figure 5.10.

[84] The beginning of the gesture is m. 3 because the celesta trill is not interrupted.
[85] Using them both for labelling is possible depending on the features of the given work.

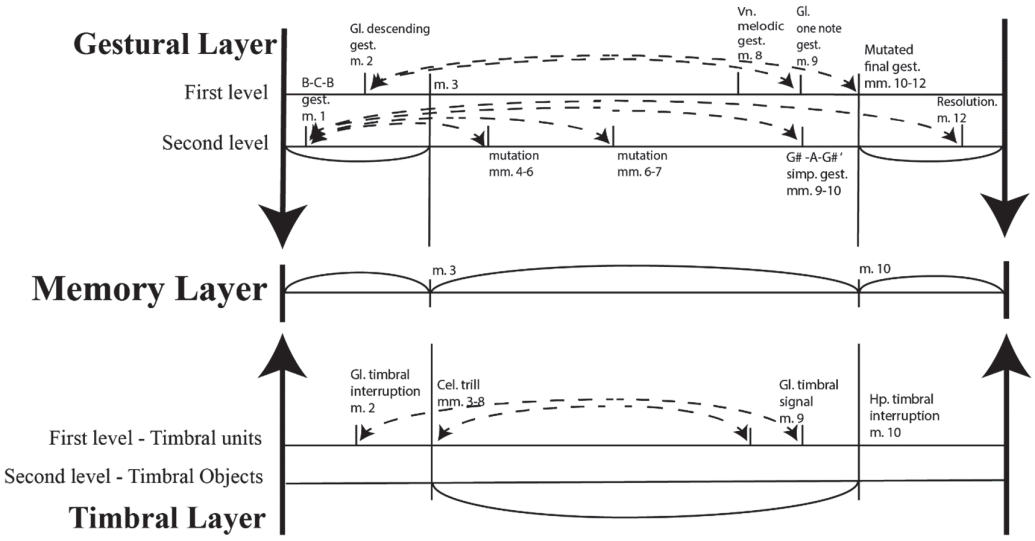

Figure 5.10 : Final form scheme for Webern's Op. 10, No. 1

Figure 5.10 shows the significances in both layers and reflects their intersectional moments into the memory layer, as it was explained in 4.6. Regions in the scheme can be used as a source to analyze the overall timbral properties and their reflection on formal timbre. Combined with the regional notation above, this scheme can be a source to discuss more topics on timbral coherence of the work.

A summary of aforementioned gestural and timbral relations can be seen in Figure 5.11.

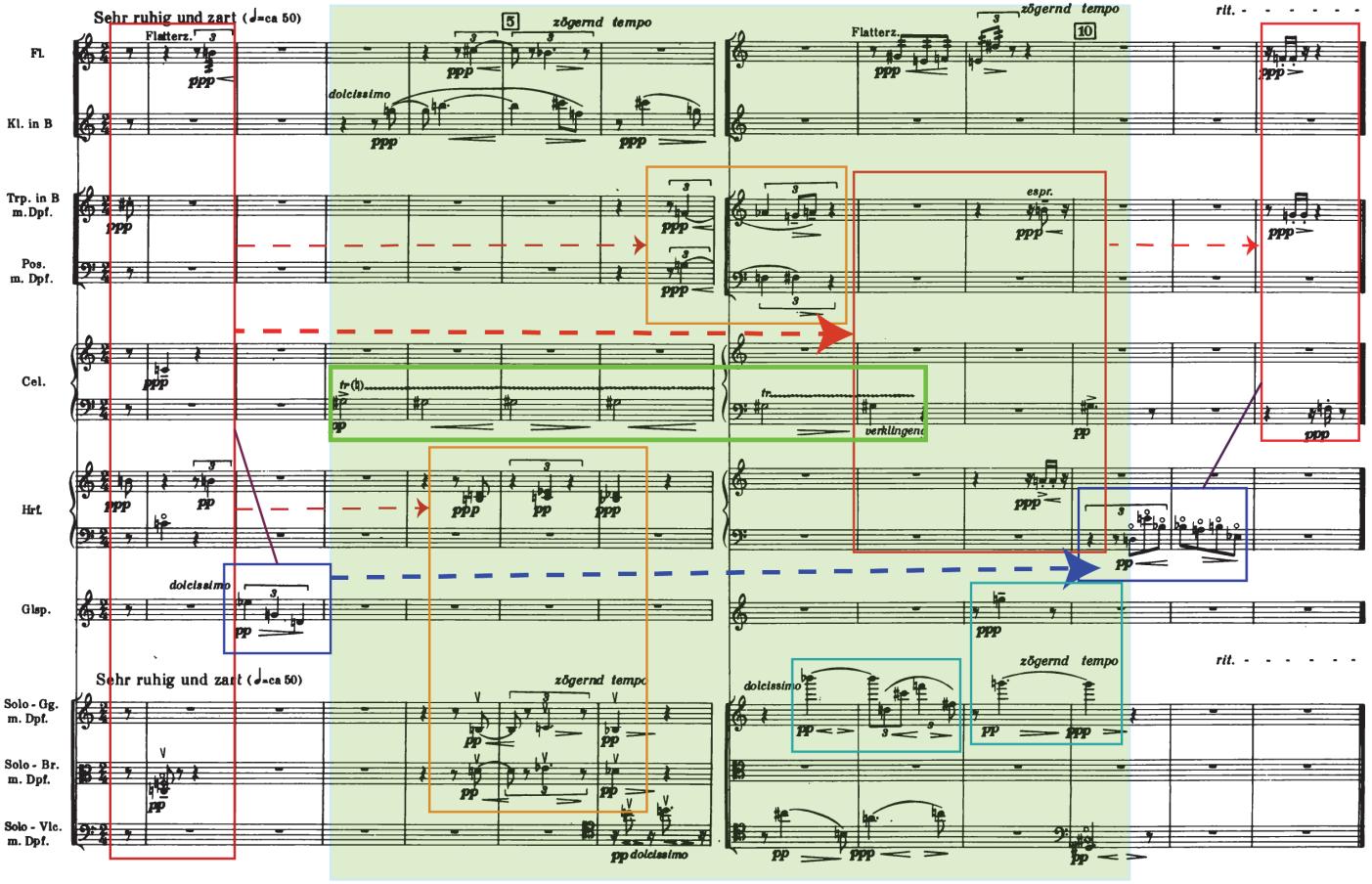

Figure 5.11 : Gestural and timbral relations presented in adapted full score.

5.2 George Crumb, *Black Angels* (1970)

George Crumb's (1929-) "Black Angels: Thirteen Images from the Dark Land (Images 1)" for "electric string quartet" was commissioned by University of Michigan and dedicated to Stanley Quartet. First page of the score includes the phrase *"in tempore belli"* [86], referring to continuing Vietnam War and last page indicates that it was "finished on Friday the Thirteenth, March 1970". Quartet was premiered on October 23, 1970; and in 1972 it was released as a recording. One movement of the piece was also used in the horror movie, The Exorcist (Tick, 2008, p. 658).

As Crumb explains, his initial intention was not his work to be interpreted as an "anti-war piece" but later he accepted that ongoing war influenced the process of composition (Keller, 2014, p. 153-154). In the liner notes for the first recording, he said that "Black Angels was conceived as a kind of parable on our troubled contemporary world." (Tick, 2008, p. 659)

The idea of 'the troubled world' is directly reflected to the visible presentation of the work, firstly with the names of the movements, and secondly with numerology. Thirteen movements are represented with dualities centered around the concepts of 'God and Devil' and the numbers seven and thirteen. These thirteen movements with the total duration of twenty minutes are grouped in three parts and their conceptual relations are indicated on a page titled 'program' in the original score.

This scheme (Figure 5.12) indeed represents the formal relations concerning the musical material gestures, furthermore the timbral features. Thus, the total form of the piece is already given by the composer. However, this is a partial representation of the form and a detailed analysis of gestural and timbral relations are needed for a final scheme.

Instrumentation of the work, as described by Crumb, has a specific set of timbral preferences. He suggests electric instruments with built-in pick-ups; or acoustic instruments with high-quality contact microphones. More importantly, he instructs to set dynamic level and reverberation of the amplifiers to be set on "high" to "create a more surrealistic effect." Performers also use maracas, tam-tams (with additional bow)

[86]Latin: "In the time of war".

and crystal glasses that are tuned in different pitches using water in order to get a "glass-harmonica effect". In this setup, all of these instruments are used as real timbral units, and not as 'imitative effects'. Other than instruments, there are certain sections in the work that have spoken words, bringing a verbal aspect.

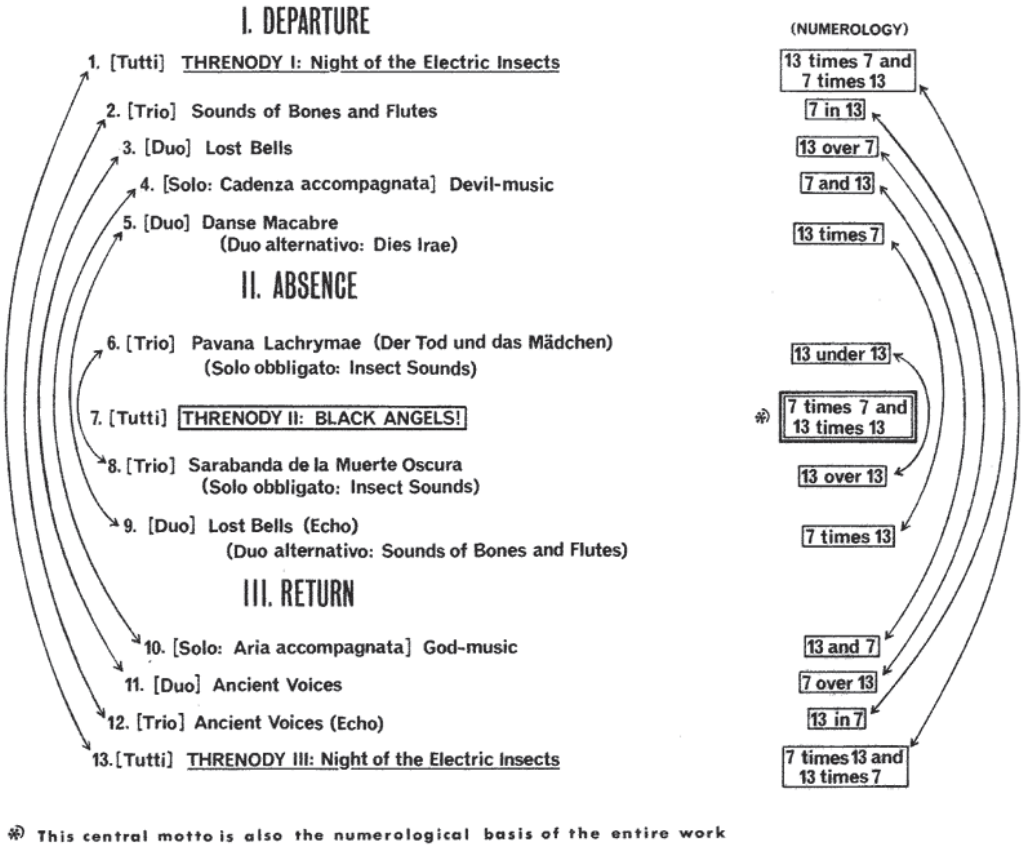

Figure 5.12 : Black Angel's program representing all movements and their conceptual and numerological relations.

A general discussion and comparison of "Threnody I: Night of the Electric Insects" is as follows.

"Threnody I: Night of the Electric Insects", starts with an extreme-high register gesture on all four string instruments. This gesture can be clearly defined with its use of tremolo and glissando (See Figure 5.13). This movement does not have any other instruments, it is highly homogenous in its spectral distribution. With the addition of the fact that the first gesture never disappears during the whole movement, contextually, the whole movement can be considered as one timbral object (See 4.4).

Here, a discussion about a melodic gesture losing its expressive identity and turning into a figure can be appropriate as it was mentioned in 3.2.2.

Timbral changes are realized through (i) addition and removal of instruments, (ii) *subito* gestures (or 'echo' effects) and (iii) dynamic gestures, *crescendo-decrescendo* continuities (or waves)[87]. Since these gestures are presented repeatedly, rather than changing the perception of timbral object, they provide another layer which belongs to the gestural domain. Therefore, for this particular movement, gestural layer includes, and directly created by the changes in timbral layer.

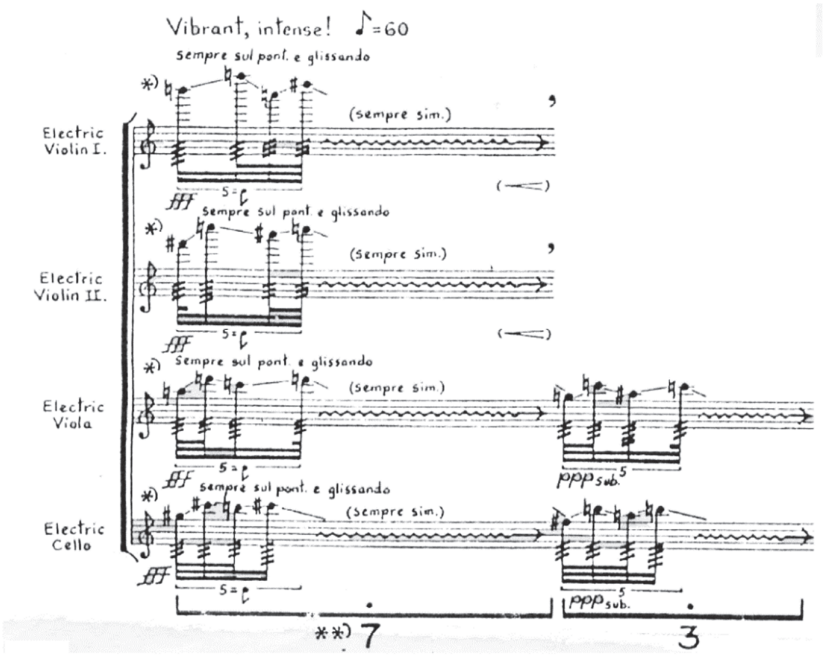

Figure 5.13 : Opening gesture of Threnody I (with its 'echo').

These dynamics changes occur mostly between ***fff*** and ***ppp***; and (ii) is always supported with (i) with the exception of one moment, where a (ii) is applied to all instruments, creating the same effect.

One last gesture that is essential to the analysis is the one that occurs twice over the ***ppp*** moments. It is a melodic gesture described with *piangendo* (iv) and it is very similar to the initial gesture with its contour and glissando effect. However, by eliminating the tremolo and adding *sul ponticello* in a higher pitch register, it creates a crucial difference in timbral coherence.

[87] All of continuous dynamic gestures are summed up in this category.

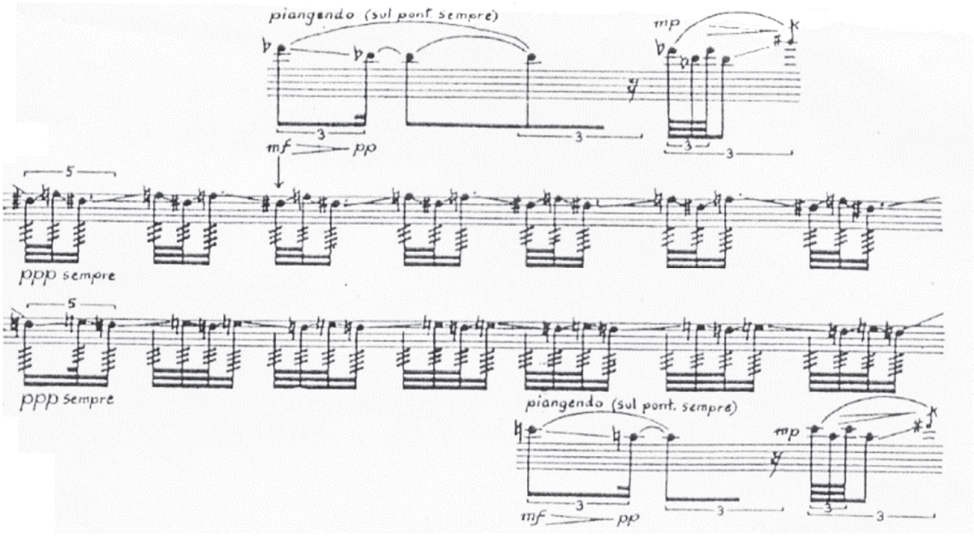

Figure 5.14 : Melodic gesture over ***ppp*** background.

Depending on the principles and analysis given above, a scheme for this movement can be concluded as follows. Despite there are durations given on score in seconds, performance[88] is not perfectly matching with the given indications.

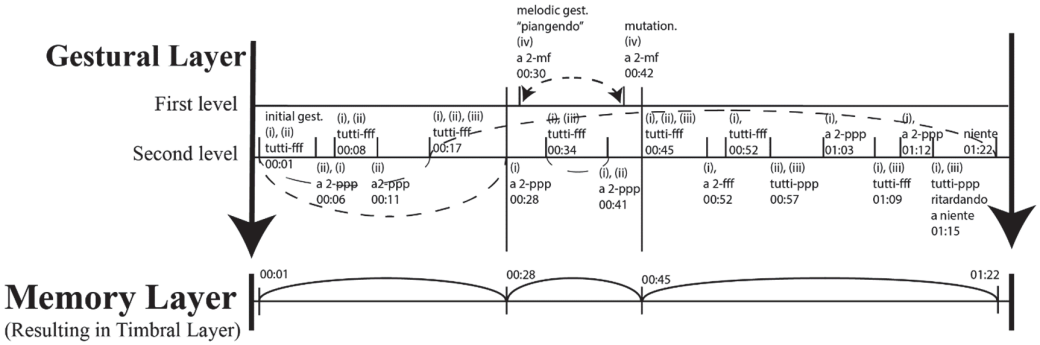

Figure 5.15 : Final form scheme for Threnody I: Night of the Electric Insects.

For practicality reasons, beginning gesture is labelled with both (i) and (ii), since it is a natural *subito* and instruments being added from nothing. Dynamic levels for each gesture entry are given due to its crucial importance for timbre.

Piangendo melodic gestures are contained by the two "a 2-***ppp***"[89] regions. These regions function as background for the melodic gesture. This is a direct contrast to pure timbral changes without any motivic material for the rest of the movement. There is

[88] Recording from Kronos Quartet's 1990 album "Black Angels" is used.
[89] "a 2" indicates that two instruments are active in that certain gesture/region. During *piangendo* gestures, they are grouped in two separate a 2's.

an overlap between the regions that contain continuous dynamic gestures (iii) and melodic gestures; but melodic region of 00:28-00:41 appears as a more significant element because it interrupts a continuous development (See 4.4).

The region after 00:45 can be interpreted as a mutation of the first region before (iv) appears. Basic material is very similar, but it uses a lower register, resulting in a high to high-mid spectrum. Another topic for debate concerning this movement can be about labelling the register and spectrum either in quartet's own feasible area, or in an absolute way that refers to total range of the orchestral instruments. This discussion is to be decided depending on the main goal of the analysis.

A possible summarized, non-detailed formal notation can be given as follows.

q=60, [(vn-1, vn-2 va, vc)(amp.;h-rev)], s. 00:01-01:22, $H_E \rightarrow H \rightarrow M_H$*
reg. "a": s. 00:01-00:28;
gest. "i"+"ii", s. 00:01-00:17;
gest. "i"+"ii"+"iii", s. 00:17-00:28;
reg. "b": s. 00:28-00:45;
gest. "i"+"iv", s. 00:28-00:34;
gest. "i"+"iii", s. 00:34-00:41;
gest. "i"+"ii"+"iv", s. 00:41-00:45;
reg. "c": gest. "i"+"ii"+"iii", s. 00:45-01:22.

In this notation, "amp" indicates that are instruments are amplified, and it is also possible to indicate certain settings like high reverb (h-rev).

Instrument names are not repeated in regions because all instruments are present in all regions. Instead, regions are directly labelled with the gestures that are present.

Gestures are notated in their order of appearance; they are not repeated in notation when they occur more than once in sub-regions.

6. CONCLUSION

As our cultural audio spaces expand, we are increasingly exposed to more genres of music and experience more modes of listening to music. Accordingly, we will have to consider how we analyze, comment on, and comprehend music; and update our methods. ‘Timbral coherence’ is an approach that is worth considering for further research, developing new methods and applications because of the increasing timbral focus in all contemporary genres of music and their listeners.

The notion of timbre and timbral coherence with its sub-topics are discussed in five chapters, starting with historical context and cognitive aspect in first two chapters. The dimensions timbre exists in, and its perceptual interactions in several aspects such as pitch, melody and dynamics are covered throughout second chapter. Questioning the relativity of timbre mostly resulted in opposite side, absoluteness. Listener seems to consider timbre an absolute quality that can stand on its own. In connection with absoluteness, liminality factor indicates that there are perceptual thresholds between timbre and pitch. Most of the time, one is barely aware of the two together. Besides, there are differences between gradual and instant transformations between pitch oriented and timbre-oriented materials. An important topic of research would be exploring these thresholds in new (and more contemporary) musical material.

Another thought coming out of sections 2.3 and 2.4 was the possibility of timbre-based grammar and its functional possibilities. Sometimes, harmonicity and inharmonicity is seen as an analogy to tonal concepts of consonance and dissonance. However, knowing more about the perceptual aspect and scientific data gathered from these studies raises doubts about this analogy and its practice in composition. As, the duration and timing topics seem more important more than before, this analogy seems a very remote one. Functional timbre-based grammar could be possible; however, these structures may have to be directed and exaggerated. Section 2.4 was an attempt to create scheme for a hierarchical organization and type-class based categorization of timbre which was based on previous conclusions of Chapter 2. These ideas are extended and functionalized in Chapter 4 and Chapter 5.

In Chapter 3, gesture, an essential dimension for all music, is examined in several aspects and categorized in order to create a functional template that could work with timbral approach. This examination is initially done in gesture's relation with language and physical gestures. Since performer's gestures are the direct source of what listener perceives, the link between physical and mental parts and versions of the gesture, and the differences between basic structures such as figure and motif are sought. These searches resulted in a multi-level categorization of gesture where it is graded in complexity and perceptual aspects. Through the perspective presented with this multi-level categorization (similar to what is done for timbre), certain examples from composer's side are examined and commented on, in order to clarify the possibilities and potential of an analytical thinking for gestures. Lastly, as a summary and an explanation of gestural functionalities, the conveyance of the musical gesture from the source to the listener is explained. The topic of conveyance, which essentially has the potential to change the perception of gesture, directly or indirectly, was essential for setting the level we are analyzing or composing.

Combining what is learned and deducted, Chapter 4 presented a proposal for a multilayered analysis of timbral and gestural networks. The basic timbral descriptors which were briefly explained in 2.4 were put into context to prepare a provisional analysis approach. At this point, the interaction and distinction between timbral objects and gestural bodies were better clarified and put into practice in Chapter 5. Analyses presented Chapter 5 were merely a beginning, a brainstorming of what would a more established method depending on the philosophical framework presented in previous chapters be. Nonetheless, the results and the insight provided by these analyses confirm the initial motivation that is presented. More principles and tools can be investigated and put into analytical applications with new works and contexts.

Depending on the research and proposed approach, the initial definition stays as a robust one. Yet, technical and practical aspects of timbral coherence approach need to be added to this definition to make it more accessible. Thus, a second -but not separate- definition would be: (In musical context) timbral coherence is an approach to analyze timbral aspects of a work in a multilayered and descriptive method in order to achieve an understanding of a perceived form, using a tailored method of gestural analysis integrated in it to set a consistent mechanism.

Further topics to be studied in order to develop the given approach would be:

- A practical and functional method to examine the complex timbres and their transpositions in degrees of congruency, in order to find possible deeper formal relations.
- Developing new experiments to test proposed principles, in order to empirically enhance the method.
- Analyzing larger scale works from different genres and periods in order to determine a basic strategy for various issues of analysis.
- Preparing specific charts for timbral features of instruments for making analysis and notation process easier.
- Embedding the approach in pedagogical context for music theory, composition, instrumentation, and orchestration.

In conclusion, timbral coherence approach can serve both as a method analysis and for the process of composition. Its deficiencies and potential should be determined in the process.

REFERENCES

Anderson, J. (2000). A provisional history of spectral music. *Contemporary Music Review*, *19* (2)*,* 7-22.

Anderson, J. (2019, June 14). Julian Anderson: Spectral centrality: proposals for the introduction of new formal harmonic structures based on discoveries in spectral composition and psycho-acoustics. [Video file]. Retrieved from https://medias.ircam.fr/xe445e2

Angerer, M. (n.d.)"Webern - 5 Stücke Op. 10 for Orchestra." *Universal Edition*, 1 Jan. 1970, https://www.universaledition.com/anton-webern-762/works/5-stuecke-op-10-745

Apperception. (1972). in Dictionary of Philosophy, Runes, D. D. (ed.), Littlefield, Adams, and Company, Totowa, NJ.

Bailey, K. (1998). *The life of Webern*. Cambridge: Cambridge University Press.

Baschet, F. (2013) Instrumental Gesture in StreicherKreis. *Contemporary Music Review*, *32* (1), 17-28.

Bausch, T. C. (2018). *The spectral atom : cohesion of spectral particles in the music of Alvin Lucier* . (Master Thesis) Electronic Theses and Dissertations. Paper 2950. https://doi.org/10.18297/etd/2950

Ben-Tal, O. (2012). Characterising Musical Gestures. *Musicae Scientiae*, *16* (3), 247-61.

Blackmore, S. (2000). *The meme machine*. Oxford: Oxford University Press.

Born, G. (2012). Music Research and Psychoacoustics. In Sterne, J. (Ed.) *The sound studies reader*. (pp. 419-426) London; New York: Routledge.

Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge: MIT Press.

Brown, J. (2014). Schoenberg and redemption. In *Cambridge University Press*.

Cadoz, C. (1988). Instrumental Gesture and Musical Composition. *ICMC 1988 - International Computer Music Conference* (pp. 1-12). *Cologne: Germany.*

Cadwallader, A. C. & David, G. (2011). *Analysis of Tonal Music: a Schenkerian Approach*. New York: Oxford University Press.

Carden, J., & Cline, T. (2019). Absolute pitch: Myths, evidence and relevance to music education and performance. Psychology of Music, *47(6)*, 890–901. https://doi.org/10.1177/0305735619856098

Clark, E. V. (2009). First language acquisition. Cambridge: Cambridge University Press.

Clarke, E. F. (1987). Levels of structure in the organization of musical time. *Contemporary Music Review*, *2* (1), 211-238.

Cogan, R. (1984). New Images of Musical Sound. Cambridge, MA: Harvard University Press.

Cone, E. T. (1974). The Composer's Voice. Berkeley: University of California Press.

Cousineau, M., Carcagno, S., Demany, L. and Pressnitzer, D. (2014). What Is a Melody? On the Relationship between Pitch and Brightness of Timbre. *Frontiers in Systems Neuroscience, 7* (127), 1-7.

Cramer, A. (2002). Schoenberg's: A Principle of Early Atonal Harmony. *Music Theory Spectrum, 24* (1)*,* 1-34*.*

Cross, I. (2007). Psychology of music, II: Perception & cognition. Retrieved December 15, 2014, Retrieved from http://0-www.oxfordmusiconline.com.divit.library.itu.edu.tr/subscriber/article/grove/music/42574pg2

Crumb, G. (1970). Black Angels: Thirteen Images from the Dark Land (Images 1). New York: Edition Peters. (1973)

Crumb, G. (1990). Black Angels [Kronos Quartet] *Black Angels* [CD]. New York: Nonesuch Records.

Delalande, F. (1988). "La gestique de Gould". In: Ghyslaine Guertin (ed.), *Glenn Gould Pluriel* (pp. 85-111). Verdun, Qc: L. Courteau.

Deutsch, D. (2002). The puzzle of absolute pitch. *Current Directions in Psychological Science*, *11* (6), 200–204. https://doi.org/10.1111/1467-8721.00200

Deutsch, D. (2013). *The Psychology of Music*. By Diana Deutsch. London: Academic.

Donnadieu, S. (2007). Mental Representation of the Timbre of Complex Sounds. In *Analysis, synthesis, and perception of musical sounds: the sound of music* (pp. 272-319). New York: Springer.

Dufourt, H. (1991). *Musique, pouvoir, écriture.* Paris: Klincksieck

Emmerson, S. (1986). *The Language of electroacoustic music*. Basingstoke, Hampshire: Macmillan Press, Music Division.

Enescu, G. (1916). Piano Suite No. 3, Op.18 Pièces Impromptues. Buchaest: Editura Muzicală. (1998).

Entrainment. (2001). Dictionary of biological psychology. London: Routledge.

Feyereisen, P. (1997). The Competition between Gesture and Speech Production in Dual-Task Paradigms. *Journal of Memory and Language*, *36* (1), 13–33.

Figure. (2001). In Grove Music Online. Retrieved 11 Dec. 2019. From https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000053712

François, J. C. (1990). Fixed Timbre, Dynamic Timbre. *Perspectives of New Music, 28* (2), 112. doi:10.2307/833012

Garcia, J., Leroux, P. & Bresson, J. (2014). pOM: Linking Pen Gestures to Computer-Aided Composition Processes. In: Proceedings of 40th International Computer Music Conference, ICMC 2014 and 11th Sound and Music Computing Conference, SMC 2014.

Gregory, A. H. (1994). Timbre and auditory streaming. *Music Perception*, *12*, 161-174.

Grisey, G. (1994-1996). Vortex Temporum I, II, III: Pour Piano et Cinq Instruments. Paris: Editions Ricordi (1995).

Hall, M. & Beauchamp, J. (2009). Clarifying spectral and temporal dimensions of musical instrument timbre. *Canadian Acoustics - Acoustique Canadienne*, *37* (1), 3-22.

Hatten, R. S. (1994). *Musical meaning in Beethoven: markedness, correlation, and interpretation*. Bloomington IN: Indiana University Press.

Hatten, R. S. (2004). *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington: Indiana University Press.

Haviland, J. B. (2001). "Gesture" In Key Terms in Alessandro Duranti (Ed.) *Language and Culture*. Malden, Massachusetts: Blackwell Publishers.

Huron, D. (1991). Tonal Consonance versus Tonal Fusion in Polyphonic Sonorities. *Music Perception: An Interdisciplinary Journal*, *9* (2), 135–154.

Huron, D. (2001). Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles. *Music Perception, 19* (1), 1–64.

Ihde, D. (2007). *Listening and voice: Phenomenologies of sound*. Albany: State University of New York Press.

Iverson, P. (1995). Auditory Stream Segregation by Musical Timbre: Effects of Static and Dynamic Acoustic Attributes. *Journal of Experimental Psychology: Human Perception and Performance, 21* (4), 751-63

Jackendoff, R., and F. Lerdahl. (2006). The Capacity For Music: What Is It, And What's Special About It? *Cognition, 100,* 33-72.

Johnson, J. (2006). *Webern and the transformation of nature*. New York: Cambridge University Press.

Keller, J. M. (2014). *Chamber music: a listener's guide*. New York: Oxford University Press.

Kendon, A. (2016). Reflections on the 'Gesture-First' Hypothesis of Language Origins. *Psychonomic Bulletin & Review, 24* (1), 163-70. Springer Link.

Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In S. Nielze´n, & O. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43-53). Amsterdam: Excerpta Medica.

Krumhansl, C. L. and Iverson P., (1992). Perceptual Interactions between Musical Pitch and Timbre. *Journal of Experimental Psychology: Human Perception and Performance, 18* (3), 739-51.

Lai, Y., Ma, W., Duan, S., & Yao, D. (2011). Involuntary Processing of Timbre: A Mismatch Negativity Study. *International Journal of Bioelectromagnetism*, *12* (4), 212-216.

Lange, K. and Czernochowski, D. (2013). Does This Sound Familiar? Effects of Timbre Change on Episodic Retrieval of Novel Melodies. *Acta Psychologica, 143* (1)*,* 136-45.

Lavengood M. L. (2017). A New Approach to the Analysis of Timbre (Ph.D. Dissertation) Available from CUNY Academic Works https://academicworks.cuny.edu/gc_etds/2188/

Lerdahl, F. (1987). Timbral hierarchies. Contemporary Music Review, *2* (1), 135–160.

Margulis, E. H. and Levine, W. H. (2006). Timbre Priming Effects and Expectation in Melody. *Journal of New Music Research, 35* (2)*,* 175-82.

McAdams, S. (1999). Perspectives on the Contribution of Timbre to Musical Structure. *Computer Music Journal, 23* (3), 85-102.

McAdams, S. (2013). Musical Timbre Perception. In Diana Deutsch (Ed.) *The psychology of music* (3rd ed., pp. 35-62). London: Academic.

McAdams, S., & Cunible, J. (1992). Perception of Timbral Analogies. *Philosophical Transactions of the Royal Society B: Biological Sciences, 336* (1278) 383-389.

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

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535* (7613), 547-550.

Merker, B. (2000). Synchronous Chorusing and the Origins of Music. *Musicae Scientiae*, 3(1 suppl).

Meyer, M., Baumann, S., & Jancke, L. (2006). Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *NeuroImage*, *32* (4), 1510-1523.

Mithen, S. J. (2007). *The singing Neanderthals: The origins of music, language, mind, and body*. Harvard University Press.

Morley, I. (2013) "Vocal Control and Corporeal control, - Vocalization, Gesture, Rhythm, Movement, and Emotion." In The Prehistory of music. Oxford: Oxford University Press.

Moscovich, V. (1997). French Spectral Music: An Introduction, *Tempo, New Series, 200*, 21-27.

Murail, T. (1982) Désintégrations, Pour 17 Instruments et Bande Magnétique. Paris: Editions Henry Lemoine (2004).

Paraskeva, S., and McAdams, S. (1997). Influence of timbre, presence/absence of tonal hierarchy and musical training on the perception of tension/relaxation schemas of musical phrases. In *Proceedings of the 1997 International Computer Music Conference, Thessaloniki* (pp. 438-441). San Francisco, CA: International Computer Music Association.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6* (7)*,* 674–681.

Porubanova M, Shaw DJ, McKay R, Xygalatas D (2014). Memory for Expectation-Violating Concepts: The Effects of Agents and Cultural Familiarity. *PLoS ONE 9* (4): e90684. doi:10.1371/journal.pone.0090684

Pressnitzer, D., & McAdams, S. (2000). Acoustics, psychoacoustics and spectral music. *Contemporary Music Review*, *19* (2), 33-59.

Radvansky, G. A., Fleming, K. J. and Simmons, J. A. (1995). Timbre Reliance in Nonmusicians' and Musicians' Memory for Melodies." *Music Perception: An Interdisciplinary Journal, 13* (2),127-40.

Rebel, J. F. (1737). Les Elémens (for orchestra). Paris: Le Clerc, et al. (n.d.)

Russo, F. A. and Thompson W. F. (2005). An Interval Size Illusion: The Influence of Timbre on the Perceived Size of Melodic Intervals. *Perception & Psychophysics, 67* (4), 559-68.

Salwiczek, L. H., and Wickler, W. (2004). Birdsong: An evolutionary parallel to human language. *Semiotica*, *2004* (151).

Sandell, G. (1998). Macrotimbre: Contribution of attack, steady state, and verbal attributes. *Journal of The Acoustical Society of America*. *103* (2966).

Schachter, C. (1999). *Unfoldings: Essays in Schenkerian Theory and Analysis.* Joseph Nathan Straus, ed. New York - Oxford: Oxford University Press.

Schenker, H. (1977). *Free Composition (Der Freie Satz).* Ernst Oster, trans. and ed. Hillsdale, NY: Pendragon Press.

Schindler, A. F. (1996). *Beethoven As I Knew Him*. Donald W. MacArdle. ed., Constance S. Jolly. trans. Mineola, New York: Dover Publications.

Schoenberg, A., & Carter, R. E. *(*1978). *Theory of harmony.* Berkeley, CA: University of California.

Scruton, R. (1997). *The Aesthetics of Music*. Oxford: Clarendon.

Scruton, R. (2009). *Understanding Music, Philosophy and Interpretation*. London: Continuum.

Scruton, R. (2014). Music and Cognitive Science. *Royal Institute of Philosophy Supplement Roy. Inst. Philos. Suppl.*: 231-47.

Siedenburg, K., Saitis, C., McAdams, S., Popper, A. N., & Fay, R. R. (2019). *Timbre: acoustics, perception, and cognition*. Cham, Switzerland: ASA Press.

Snyder, B. (2001). *Music and Memory: An Introduction*. Cambridge, MA: CogNet.

Sonic Visualiser (Version 3.3) [Computer software] Chris Cannam, Christian Landone, and Mark Sandler. London: London: Centre for Digital Music at Queen Mary, University of London.

Stockhausen, K., & Barkin, E. (1962). The Concept of Unity in Electronic Music. Perspectives of New Music, 1(1), 39.

Szász, T. (2011). Traditional and Oriental models of Enesco's bell-inspired piano music: Sonate F+ minor (1912-24) / Choral - Carillon nocturne (1916) in *Symposium: Zwischen Zeiten & Welten Leben und Werk von Béla Bartók und George Enescu* 18-20 November, 2011, Delmenhorst.

Teodorescu-Ciocanea, L. (2003). Timbre versus spectralism. *Contemporary Music Review*, *22* (1- 2), 87-104.

Terhardt, E. (1974). Pitch of pure tones: Its relation to intensity. In E. Zwicker & E. Terhardt (Eds.), *Facts and models in hearing* (pp. 353–360). Berlin/Heidelberg: Springer.

Tick, J. (2008). *Music in the Usa: a documentary companion*. Oxford: Oxford University Press.

Timbre. (n.d.). In Lexico online dictionary Retrieved February 20, 2018, from https://www.lexico.com/

Timbre. (n.d.). In Merriam-Webster's online dictionary. Retrieved February 20, 2018, from https://www.merriam-webster.com/dictionary/

Toop, R. (1979). Stockhausen and the sine-wave: The story of an ambiguous relationship. *Musical Quarterly*, *65* (3), 379–391.

Url-1: Anton Webern (1883-1945) Fünf Stücke für orchester, op. 10 (1911-1913) on Ressources.ircam. Retrieved from http://brahms.ircam.fr/works/work/12636/

Url-2: Webern, Anton, 1883-1945. *Stücke, orchestra, op. 10* . Five pieces for orchestra, op. 10: autograph manuscript, 1913?, (n.d) on The Morgan Library & Museum, Retrieved from http://www.themorgan.org/music/manuscript/115913

Vines, B.W., Wanderley M.M. ,Krumhansl C.L. , Nuzzo R.L. , and Levitin D.J. . (2003). "Performance Gestures of Musicians: What Structural and Emotional Information Do They Convey?" *Gesture-Based Communication in Human-Computer Interaction Lecture Notes in Computer Science*: 468-78.

Vinoo, A., and Toiviainen, P. (2012). Effect of Enculturation on the Semantic and Acoustic Correlates of Polyphonic Timbre. *Music Perception: An Interdisciplinary Journal, 29* (3), 297-310.

Vurma, A., Raju, M. and Kuuda, A. (2011). Does Timbre Affect Pitch?: Estimations by Musicians and Non-musicians. *Psychology of Music 39* (3), 291-306.

Wanderley M.M., Vines, B.W., Middleton N., Mckay, C. and Hatch W. (2005). The Musical Significance of Clarinetists' Ancillary Gestures: An Exploration of the Field. *Journal of New Music Research, 34* (1), 97-113.

Webern, A. (1913). 5 Stücke für Orchester, Op.1. Vienna: Universal Edition. (1923, reprint 1951).

Webern, A. (2000). 5 Pieces for Orchestra, Op. 10 [Ensemble Intercontemporain, Pierre Boulez] Complete Works [CD] Germany: Deutsche Grammophon.

Wessel, D. L. (1979). Timbre Space as a Musical Control Structure." *Computer Music Journal, 3* (2), 45-52.

Willoughby, H. (2000). The Sound of Han: Pansori, Timbre and a Korean Ethos of Pain and Suffering. *Yearbook for Traditional Music*, 32, 17-30.

Winn, P. (2001). Dictionary of biological psychology. London: Routledge.

Wolpert, R. (1990). Recognition of Melody, Harmonic Accompaniment, and Instrumentation: Musicians vs. Nonmusicians. *Music Perception: An Interdisciplinary Journal*, *8 (1),* 95-105.

Wong, Y. K., Lui, K. F. H., Yip, K. H. M., & Wong, A. C. N. (2019). Is it impossible to acquire absolute pitch in adulthood? *Attention, Perception, and Psychophysics.* Advance online publication https://doi.org/10.3758/s13414-019-01869-3

Word Painting. (2001). "Word-painting." In *Grove Music Online*. Retrieved 10 November 2019, from https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000030568

Wright, J. K., & Bregman, A. S. (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review, 2* (1), 63-92.

Zbikowski, L. M. (2011). Musical Gesture and Musical Grammar: A Cognitive Approach. In *New Perspectives on Music and Gesture*. Ed. Anthony Gritten and Elaine King. Farnham, Surrey: Ashgate Publishing Ltd.

APPENDICES

APPENDIX A: Opening pages of Jean-Féry Rebel's *Les Élémens*
APPENDIX B: Abbreviations for Musical Instruments
APPENDIX C: Musical Instrument Range Chart
APPENDIX D: Score of Webern, Op. 10, *Fünf Stücke für Orchester*, No. 1 (1913)

APPENDIX A: Opening pages of Jean-Féry Rebel's *Les Élémens* (1737)

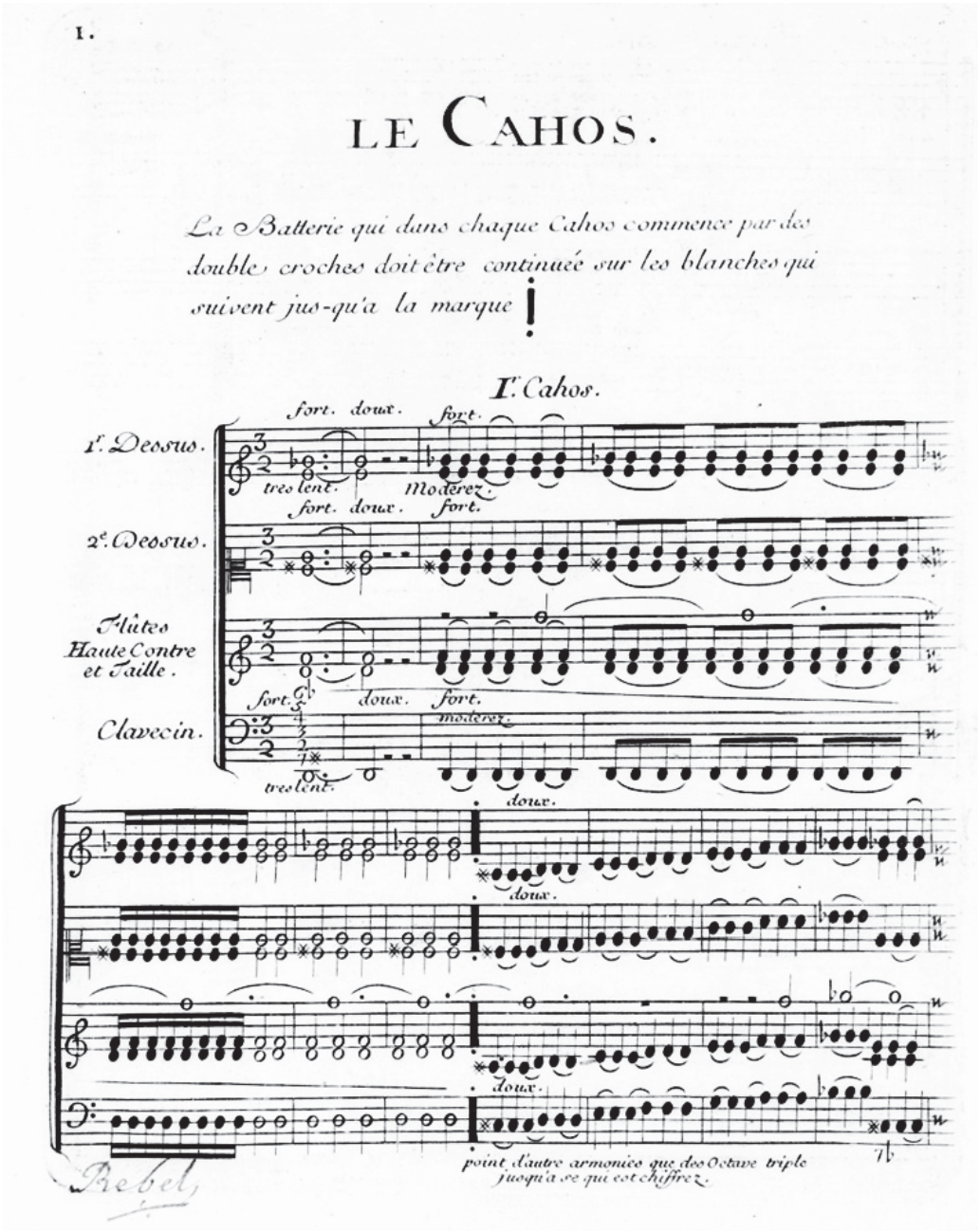

Figure A.1 : Full score of *Les Élémens.*

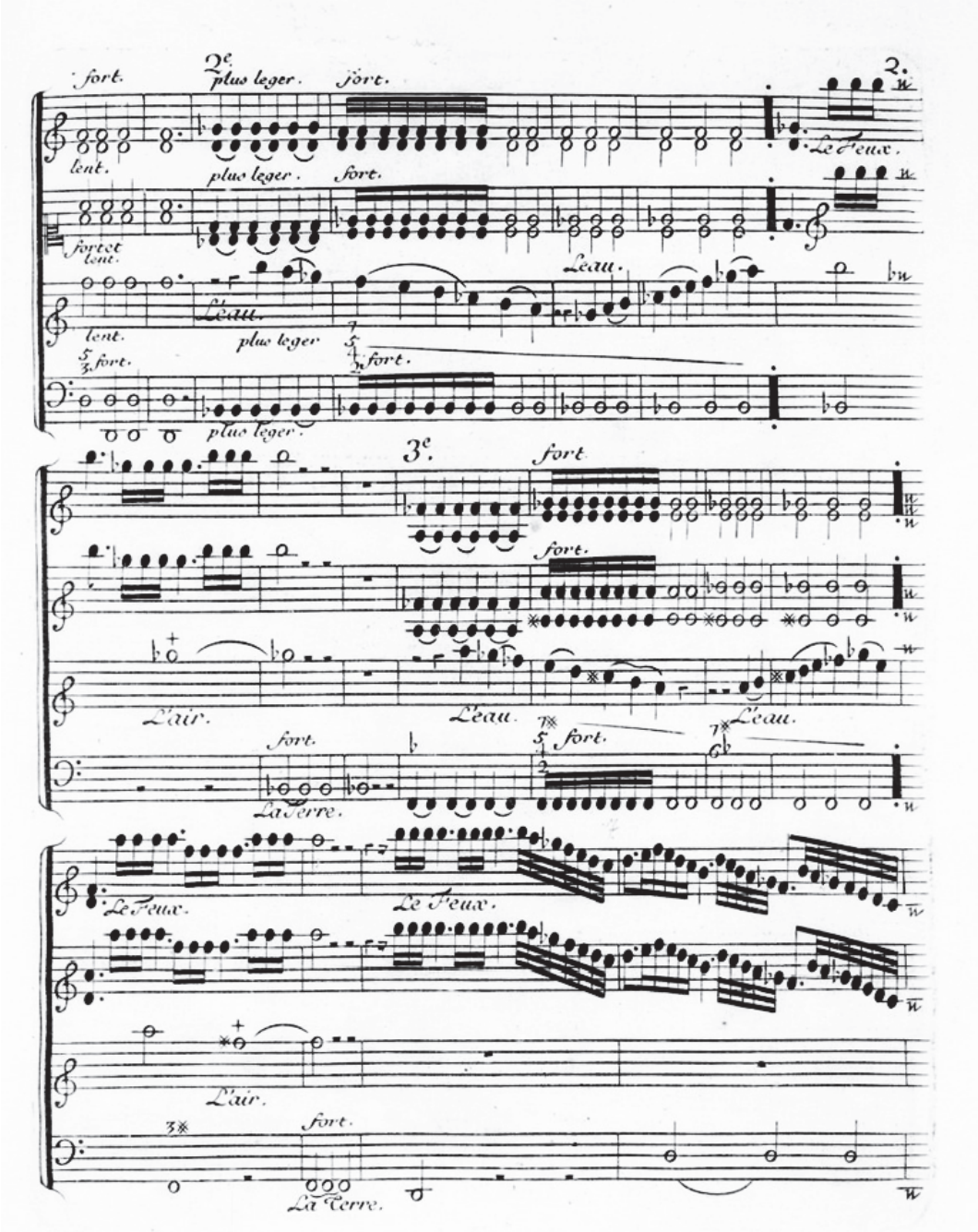

Figure A.1 (continued) : Full score of *Les Élémens.*

APPENDIX B: Abbreviations for Musical Instruments

Abbreviation	Instrument
acc	Accordion
afl	Alto flute
alt	Alto (voice) (contralto)
arp	Arpeggione
bag	Bagpipe
bar	Baritone (voice)
bass	Bass (voice)
bbar	Bass baritone (voice)
bc	Continuo (Basso continuo)
bcl	Bass clarinet
bell	Bell (Chimes)
bfl	Bass flute
bgtr	Bass guitar
bjo	Banjo
bn	Bassoon
bob	Bass oboe (Baritone oboe)
br	Brass instruments
bryt	Baryton
bstcl	Basset clarinet
bsthn	Basset horn
bug	Bugle
cbcl	Contrabass clarinet
cbn	Contrabassoon
cch	Children's chorus
cel	Celesta
ch	Mixed chorus
cimb	Cimbalom
cit	Cittern
cl	Clarinet
clvd	Clavichord
cm	Chalumeau
conc	Concertina
crh	Crumhorn
crt	Cornet
crtt	Cornett (Zink)
cv	Child's voice
db	Double Bass
dlcn	Dulcian
dom	Domra
dulc	Dulcimer
egtr	Electric guitar
eh	English horn (Cor anglais)

Abbreviation	Instrument
elec	Electronic Instruments
epf	Electric piano
eq	Equal voices
erhu	Erhu
euph	Euphonium
fch	Female chorus
fda	Flute d'amore (Tenor flute)
fgh	Flugelhorn
fife	Fife
fl	Flute
flag	Flageolet
ghca	Glass harmonica (Bowl organ)
gl	Glockenspiel
gtr	Guitar
harm	Harmonium
hca	Harmonica (Mouth Organ)
heck	Heckelphone
hn	Horn
hp	Harp
hpd	Harpsichord
kbd	Keyboard instrument
lute	Lute
lyre	Lyre
mand	Mandolin
mar	Marimba
mch	Male chorus
mez	Mezzo-soprano
mus	Musette
nar	Narrator (Reciter)
ob	Oboe
oca	Ocarina
oda	Oboe d'amore
om	Ondes Martenot
oph	Ophicleide
orch	Orchestra
org	Organ
oud	Oud
pan	Pan flute (Pan-pipes)
perc	Percussion
pf	Piano
pf3h	Piano 3 hands
pf4h	Piano 4 hands

Figure B.1 : Abbreviations for Musical Instruments (Retrieved from Petrucci Music Library https://imslp.org).

Abbreviation	Instrument
pf5h	Piano 5 hands
pf6h	Piano 6 hands
pflh	Piano left hand
pfped	Pedal piano
pfrh	Piano right hand
picc	Piccolo
pipa	Pipa
pk	Timpani
ptpt	Piccolo trumpet
reb	Rebec
rec	Recorder
sar	Sarrusophone
sax	Saxophone
sheng	Sheng
shw	Shawm
sit	Sitar
skbt	Sackbut
sop	Soprano (voice)
srp	Serpent
stpt	Slide trumpet
str	String instruments
sxh	Saxhorn
syn	Synthesizer
tba	Tuba

Abbreviation	Instrument
tbn	Trombone
ten	Tenor
thrm	Theremin
timp	Timpani
tpt	Trumpet
uch	Unison chorus
uke	Ukelele (Ukulele)
v	Voice (solo)
va	Viola
vap	Viola pomposa
vc	Cello
vda	Viola d'amore
vib	Vibraphone
vie	Vielle (Hurdy-Gurdy)
viol	Viol (Viola da gamba)
vlne	Violone
vn	Violin
vuv	Vuvuzela
vv	Voices (multiple soloists)
wag	Wagner tuba
ww	Woodwind instruments
xiao	Xiao
xyl	Xylophone
zith	Zither

Figure B.1 (continued) : Abbreviations for Musical Instruments.

APPENDIX C: Musical Instrument Range Chart

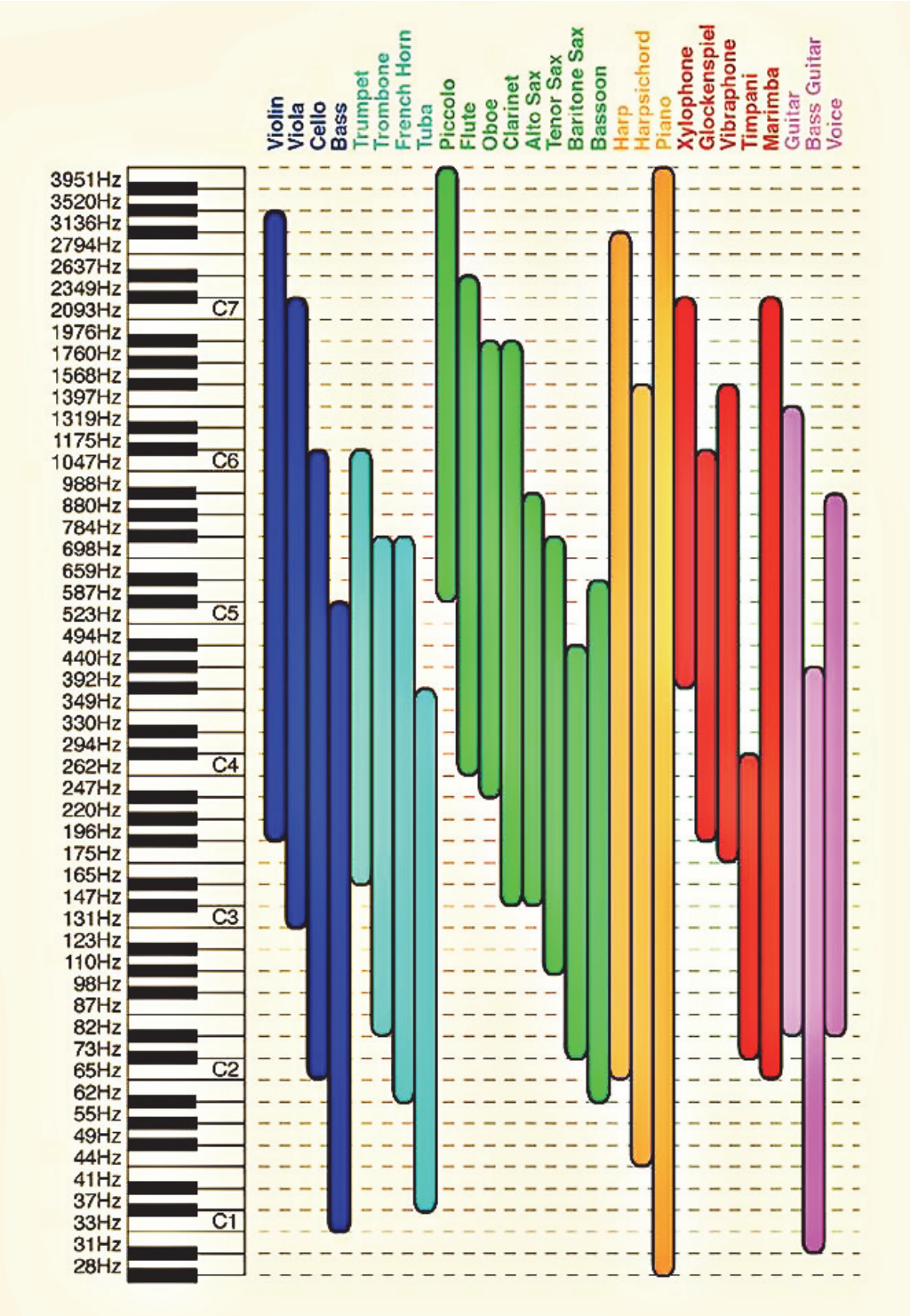

Figure C.1 : Musical Instrument Range Chart (Wikipedia image under the Creative Commons Attribution-Share Alike 3.0 Unported license).

APPENDIX D: Score of Webern, Op. 10, *Fünf Stücke für Orchester*, No. 1 (1913)

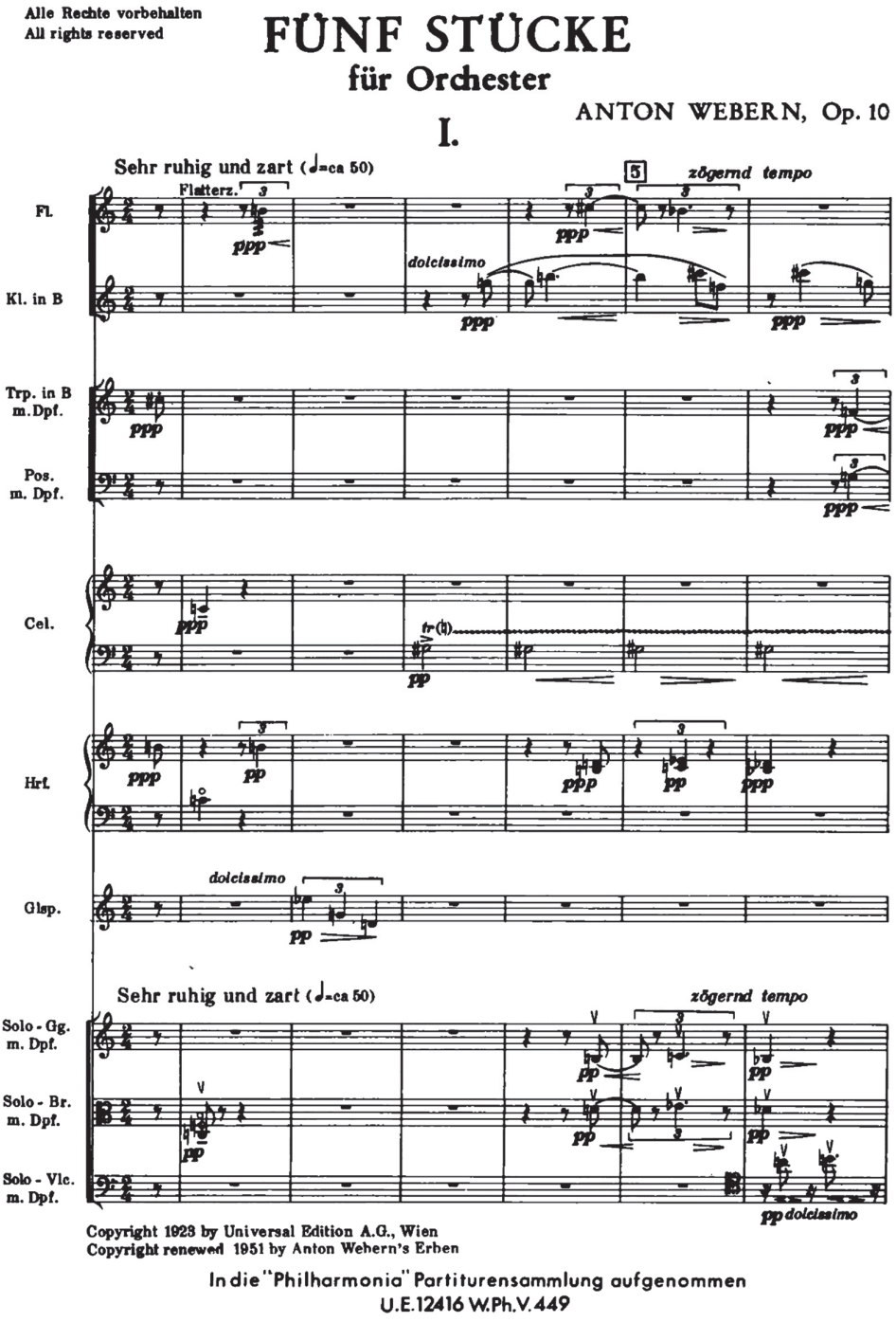

Figure D.1 : Full score of Webern, Op. 10, No. 1.

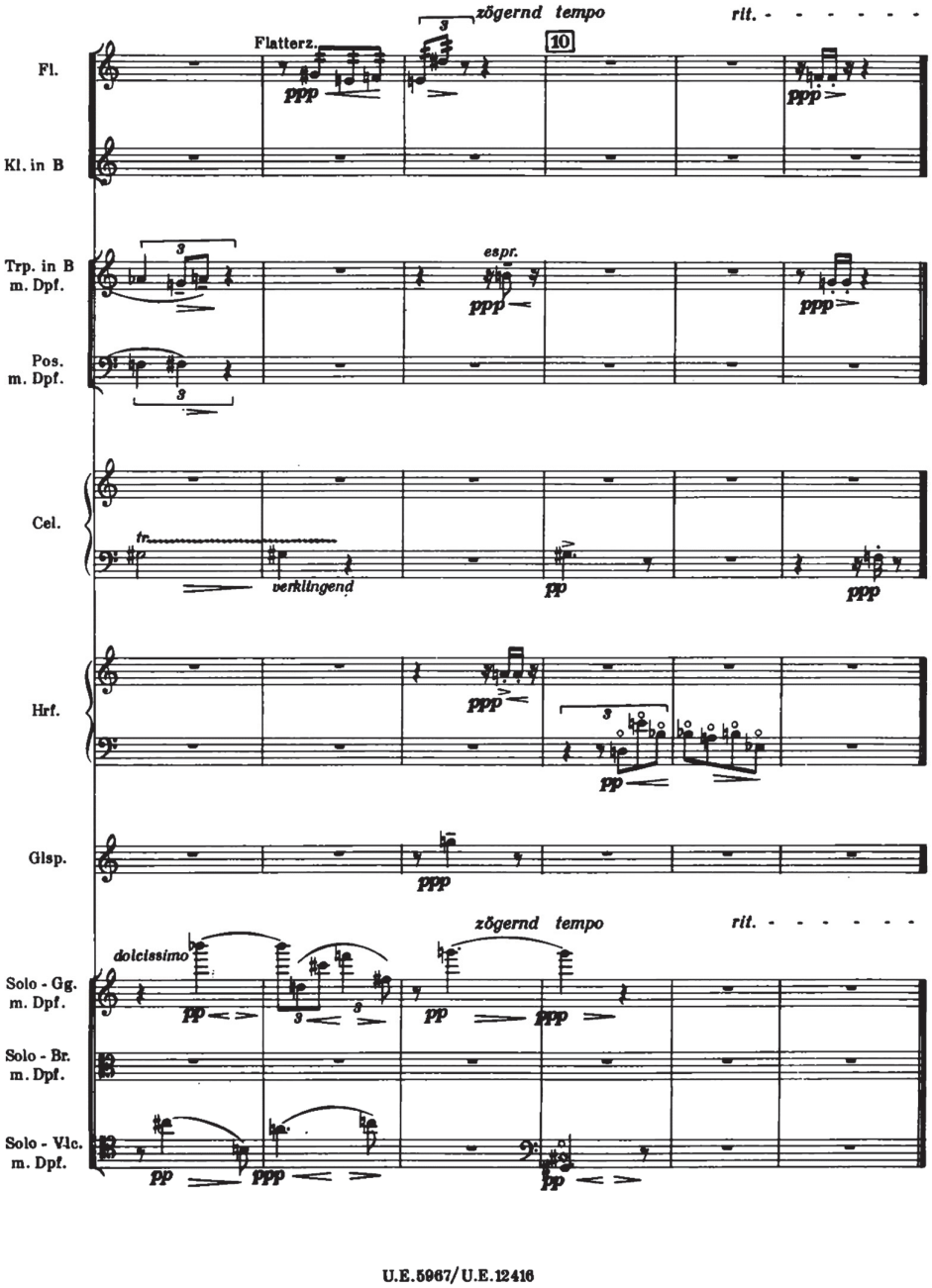

Figure D.1 (continued) : Full score of Webern, Op. 10, No. 1.

CURRICULUM VITAE

Name Surname **:** Enis Gümüş

Place and Date of Birth **:** Samsun, 21 June 1982

E-Mail **:** enis.gumus@gmail.com

EDUCATION **:**

- **B.A.** **:** 2010, Yıldız Technical University, Faculty of Art and Design, Department of Music and Performing Arts, Audio Design Programme
- **M.Mus.** **:** 2012, Codarts University of the Arts, Rotterdam Conservatory, Master of Music Programme in Classical Composition

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2016 - Full-Time Lecturer, Department of Music, İstanbul Bilgi University
- 2014 - Independent Music Professional, Composer, Sound Designer, Performer, Teacher; Istanbul
- 2015 - 2016 Research Assistant, Department of Music, İstanbul Bilgi University
- 2011 - 2013 Conductor, Teacher, Rotterdam Turkish Classical Music Choir
- 2009 - 2013 Independent Music Professional, Composer, Teacher; Rotterdam

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Gümüş, E.** 2019. Gestural Hierarchies and Their Implications in Compositional Structures, *Porte Akademik*, *18-19*, 71-85.