THE ROLE OF EXPECTATION IN THE CONSTITUTION
OF SUBJECTIVE MUSICAL EXPERIENCE

Direttore della Scuola : Ch.mo Prof. Giovanni Fiaschi
Coordinatore d'indirizzo: Ch.ma Prof.ssa Francesca Menegoni
Supervisore : Ch.mo Prof. Antonio Maria Nunziante

Dottoranda : Elisa Negretto
ABSTRACT

The present study is a theoretical discussion concerning some of the important processes that characterize human perception, which is understood as a fundamental structure of consciousness. The aim is to acquire new insights for a better comprehension of the human experience in the world and the way individual subjects become familiar with their environment.

To accomplish this task, the experience of listening to music is analysed due to the widespread acceptance of music as an important aspect of human life. With reference to various research studies and hypothesis focused on the way listeners understand music, principles of perceptual organization which allow immediate awareness of music are investigated. Particular attention is given to studies in the field of music perception and cognition.

Through an analysis of the main mental processes involved, the goal is to describe several relevant aspects of how listeners understand music at the perceptual level and constitute musical meanings in a subjective manner. This constitution results in a listener’s ability to become immediately aware of their auditory experience as having a specific meaning – that of being, first of all, a musical experience.

Focusing on everyday experience and exposure to music, expectation and anticipation are presented as fundamental principles that contribute to the knowledge of the auditory environment and the meaning which acoustic events acquire for a subject. An in-depth analysis of their specific features will show the different influence expectation and anticipation have on a listener’s basic perceptual comprehension of music and in the constitution of musical meanings that are subjectively characterized.

Within the framework of a theoretical investigation into human processes of musical knowledge acquisition, the additional objective is to create a dialogue between empirical research results in the field of music cognition and phenomenological descriptions of the structures of human consciousness. In particular, through clarification of the specific terminology that characterizes these two fields of research, the intent is to go beyond the historical gap which has existed between them. This combination of perspectives may increase our understanding of the way human beings know the world and learn how to react to the events which they encounter.
Contenuto della dissertazione

Con il presente lavoro viene proposta una indagine teoretica su alcuni importanti principi cognitivi che caratterizzano la percezione umana, considerata come una struttura fondamentale della coscienza. Lo scopo è contribuire ad una migliore comprensione dell’esperienza umana e del modo in cui singoli individui acquisiscono familiarità con l’ambiente in cui vivono.

A tal fine, e sulla base del vasto riconoscimento della musica quale importante aspetto della vita umana, viene analizzata l’esperienza dell’ascolto musicale. Con riferimento a varie ricerche empiriche e riflessioni teoriche indirizzate all’analisi del modo in cui gli ascoltatori comprendono la musica, vengono indagati i principi di organizzazione percettiva che permettono la consapevolezza immediata della musica percepita.

L’analisi mira quindi a descrivere alcuni degli aspetti che caratterizzano la comprensione percettiva della musica e il modo in cui gli ascoltatori costituiscono esperienze musicali soggettive. Tale costituzione risulta nella capacità dell’ascoltatore di essere consapevole delle sue esperienze uditive come musicalmente significative – tali cioè da essere in primo luogo esperienze musicali – in modo immediato e senza la necessità di ricorrere ad una riflessione cosciente.

Focalizzando l’attenzione sull’esperienza musicale quotidiana, ‘aspettative’ e ‘anticipazioni’ sono presentate quali principi fondamentali per la conoscenza dell’ambiente uditivo circostante e il significato che eventi acustici acquisiscono per un soggetto. Una analisi dettagliata delle loro caratteristiche specifiche mostrerà la diversa influenza che aspettative e anticipazioni hanno sulla comprensione percettiva di un evento musicale da parte dell’ascoltatore e sulla costituzione di significati musicali che sono soggettivamente caratterizzati.

All’interno di una riflessione teoretica sui processi di acquisizione di conoscenze musicali, l’ulteriore obiettivo è di creare un dialogo tra scienze cognitive e descrizioni fenomenologiche della struttura della coscienza umana. In particolare, attraverso la chiarificazione della terminologia specifica che caratterizza questi due campi di ricerca, lo scopo è superare le incomprensioni storicamente esistite tra loro. Tale combinazione di prospettive potrebbe accrescere la comprensione del modo in cui gli esseri umani conoscono il mondo e apprendono come reagire agli eventi di cui hanno esperienza.
# THE ROLE OF EXPECTATION IN THE CONSTITUTION OF SUBJECTIVE MUSICAL EXPERIENCE

## INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic musical understanding</td>
<td>22</td>
</tr>
<tr>
<td>Levinson’s notion of basic musical understanding</td>
<td>25</td>
</tr>
<tr>
<td>The perceiver–environment relationship: Clarke and the ecological approach to musical listening</td>
<td>37</td>
</tr>
<tr>
<td>The influence of learning in the perceptual organization of music and the possibility of innate perceptual structures and abilities</td>
<td>45</td>
</tr>
<tr>
<td>The subjective aspect of music comprehension</td>
<td>52</td>
</tr>
<tr>
<td>Intentionality</td>
<td>55</td>
</tr>
<tr>
<td>Basic musical understanding</td>
<td>59</td>
</tr>
</tbody>
</table>

## THE CONSTITUTION OF SUBJECTIVE MUSICAL EXPERIENCES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>62</td>
</tr>
<tr>
<td>The problem of musical meaning</td>
<td>62</td>
</tr>
<tr>
<td>The constitution of a meaningful musical experience</td>
<td>69</td>
</tr>
<tr>
<td>Introduction</td>
<td>69</td>
</tr>
<tr>
<td>The concept of ‘constitution’</td>
<td>71</td>
</tr>
<tr>
<td>Introduction to expectation</td>
<td>85</td>
</tr>
</tbody>
</table>

## THE EXPECTATION PROCESS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>90</td>
</tr>
<tr>
<td>Main theories about musical expectation</td>
<td>90</td>
</tr>
<tr>
<td>Main features of musical expectation</td>
<td>92</td>
</tr>
<tr>
<td>Empirical studies about expectation</td>
<td>98</td>
</tr>
<tr>
<td>Some reflections</td>
<td>103</td>
</tr>
</tbody>
</table>

## CONCEPTUAL DISTINCTION: EXPECTATION VERSUS ANTICIPATION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>115</td>
</tr>
<tr>
<td>Discussion</td>
<td>115</td>
</tr>
</tbody>
</table>

## THE PERCEIVED MUSICAL STRUCTURE VERSUS THE NOTATED/ESTABLISHED MUSICAL STRUCTURE

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>130</td>
</tr>
<tr>
<td>The concept of “structural listener”</td>
<td>135</td>
</tr>
<tr>
<td>Preliminary conceptual distinctions</td>
<td>136</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

This dissertation is the result of the work done not only by myself, but also of several other people who contributed directly or indirectly towards its production. For this reason I owe my gratitude to all those people who made this dissertation possible and gave me the opportunity to have this important experience – both at an academic and personal level. Although I cannot recall the names of all these contributors, they are present in the various parts that compose my work together with the passion which guided me to this point.

I am grateful first of all to my supervisor in the department of Philosophy at the University of Padua, Dott. Antonio M. Nunziante. I have been fortunate to have a supervisor who has given me the freedom to explore on my own, and at the same time the guidance to recover when my steps falter.

During these years I have had the great opportunity to develop my research at McGill University, Montréal (Canada), which has been the most insightful academic environment for the type of study I have conducted.

This would not be possible without the contribution of Prof. Eric Lewis, professor of philosophy at McGill, to whom I address my deepest gratitude. His comments and constructive criticisms at different stages of my research were thought-provoking and they helped me focus and enrich my ideas.

I am particularly grateful to Prof. Stephen McAdams, professor at the department of Music Theory at Schulich School of Music (McGill University), for the valuable discussions that helped me better understand my research area and create that interdisciplinary dialogue which is the aim of my work.

I would like to acknowledge Prof. Ian Gold, department of philosophy at McGill, for his encouragement and practical advice. He was always there to listen and assist with the structure of my dissertation.

I am also thankful to Prof. Philip Buckley, department of philosophy at McGill, for the discussions that helped me to better clarify and understand the phenomenological perspective from which I developed the main topics of my dissertation.

A special thank goes to my family, who constantly encouraged me throughout these years. Most importantly, none of this would have been possible without the love and patience of Paolo, who has been a constant source of strength. I deeply appreciate his belief in me.

Finally, I greatly valued the friendships and support of my colleagues, with whom I shared unforgettable experiences.
INTRODUCTION

Auditory perception reveals a paradox similar to that encountered in all realms of perception: nothing seems more simple than to perceive the sound environment around us, and yet it is a phenomenon that appears to rebel against scientific analysis. What difficulty exists, for example, in recognizing one's name in a sentence pronounced by another person, in differentiating the noise of a car from that of an airplane, in perceiving the captivating rock rhythms of Bill Haley, in recognizing the voice of one's own child or the steps of a familiar person? All that seems to be needed is simply to ‘open one’s ears’.  

In their everyday lives, people have a lot of very different auditory experiences, which are understood in different ways and acquire specific meanings in relation to their subjective experiences and ways of perceiving the living world. Music is part of the life of all people and it plays a great role in every culture as a means of social communication. It has been argued that the capacity to create and perceive music is universal, and finds its origin in specific mechanisms of the human brain. Music is so common and easily experienced without any kind of training (children, for example, are immediately able to listen to music and show interest and involvement with it) that, in order to listen to and enjoy music, “all that seems to be needed is simply to ‘open one’s ears’”.

But the phenomenon is much more complex than it seems at first sight, as it involves important perceptual processes and exposure-acquired knowledge. The various factors and processes involved in music perception might be of particular interest in an attempt to better understand the human experience in the living world and how individual subjects become familiar with it.

Consider, for example, the case of a man who, in his daily experience, gets off of the metro and identifies some sounds as music without seeing where they come from. He does not expect to listen to music in that context and he does not have the intention of

---

1 McAdams & Bigand, 1993, p. V.
2 For an overview, see Peretz & Zatorre (2003); McDermott & Hauser (2005); Wallin, Merker & Brown (2001).
listening to music as he might, for example, when going to a concert. In a metro station, the acoustical environment is very complex and it is not easy to perceive music. In spite of this, the man of our example is able to distinguish and be aware of music in the midst of other auditory events (the noise of the metro, people who are speaking, children who are screaming). He is also able to distinguish details about the music such as whether it involves someone singing, playing the violin or the guitar. Thus, he has a specific comprehension of an unexpected auditory experience. He is able to perceive music and finer details such as instrumentation instantaneously, without any kind of reflection or the mediation of a conscious thought.

A problem arising from this situation is: how is the man in the metro able to perceive music and to be aware of his subjective experience in such a complex auditory environment, without the mediation of an act of reflection?

According to Roger Scruton, the man perceptually acts a passage from an ‘acoustical experience’ to a ‘musical experience of sounds’. The problem he wishes to address is how such a passage is realized and in which way the man ‘opens his ears’ toward the music a musician is playing somewhere in the metro station. The attempt of the work is therefore to understand, at least in some of its aspects, his aptitude in discriminating music from other sounds occurring at the same time and constituting a subjective musical experience.

In relation to this problem, different questions need to be addressed. In particular: what is constitutive to the musical experience in the metro? How does the man in the metro understand and identify sounds as constituting music? These two questions are related to two complementary aspects of the problem. The first one concerns the features of music that the man picks up at the auditory level which allow for the structuring of sound events into musical forms. Therefore, it concerns the acoustic features of sound which determine it to be a musical sound (and, thus, distinguish it from the voice of a woman walking next to him or the noise of the metro train leaving the station).

There exists a large body of literature addressing this problem, and thus we are able to outline and discuss many different theories and related historical debates. For example,
Helmholtz’s acoustic theory (1967) supports the idea that musical sounds have particular acoustic features that distinguish them from all other kinds of sounds (noises), and that the human auditory system is particularly suited to recognize such sounds as music.

The second aspect of the problem that I have highlighted concerns the cognitive processes involved in the organization of sounds into musical forms. In relation to our example, this pertains to what the man in the metro mentally does to hear music. More precisely, it concerns the problem of what he perceptually does in order to distinguish music from all the other sounds and auditory events that he experiences in the context of the metro station. With the present study, I focus my attention on this aspect of the problem and, from a phenomenological perspective, I will address three main questions.

Husserl’s interpretation of time-consciousness is an attempt to reply to the following question: “how, in a flow of consciousness, is the awareness of a temporally extended object constituted?” (Brough, 2005, p. 248). Based on this problem and with reference to the case of the man in the metro, my first question is: how is the awareness of a musical event – a temporally extended object – constituted by the man among the complex auditory environment of the metro?

In Emmanuel Bigand’s words, and in relation to the complex acoustical and temporal nature of musical stimuli, the question concerns “how listeners manage to transform these vibrations into a set of sound signals having specific auditory qualities and coherence” (1993, p. 231). Additionally, in relation to the complex sequential organization of music, how do listeners manage to perceive relationships between sound events separated in time? Focusing on the way such an experience is constituted, a related question is, according to Bigand, “what mechanisms of information processing are used by listeners to attribute a particularly rich and expressive symbolic value to what was originally just a series of acoustic signals?” (1993, p. 233).

The case of the man in the metro is useful in attempting to understand and describe how listeners meaningfully represent what they hear and how they become aware of a constituted musical experience. The simple fact that the man is capable of classifying what he hears as ‘music’ indicates a kind of comprehension of the acoustical environment that
presupposes a specific organization of the information. From this perspective, the perception of music represents a particular kind of organization which allows the constitution of musical meanings and the awareness of music in a complex auditory environment. My aim is to understand, at least in some of its relevant aspects, how such organization is realized.

With particular attention paid to the development of the cognitive psychology of music, and in the attempt to relate research results in this field to a phenomenological description of the human perceptual structures of consciousness, my second question is: what are the main mental processes that determine the subjective character of the man’s musical experience in the metro? This question requires an analysis of various empirical studies and hypotheses considering what listeners mentally do to hear music, which perceptual structures are involved and which cognitive principles are used. Thus, I focus on the way listeners organize and understand sounds as music at the perceptual level in order to find the key elements that determine musical experience.

Finally, I pay particular attention to the specific case in which listeners encounter unfamiliar music. Therefore, the third question is: what mental processes are involved in the specific case where the man in the metro hears music which is unfamiliar to him?

Through the analysis and discussion of these issues I will attempt to create a dialogue between empirical research results in the field of music cognition and phenomenology within the framework of a theoretical investigation about the human processes of knowledge acquisition. As McAdams and Bigand observe,

Etymologically speaking, the term ‘cognition’ refers to the notion of knowledge. It has been used in a more specific sense to designate the conditions that allow humans to develop knowledge of the world. It almost goes without saying that no knowledge can be acquired in the absence of perceiving: in other words, no theory of knowledge is complete without a theory of its acquisition, and thus of perception. To emphasize the cognitive aspects of audition is thus primarily to remind us that auditory information participates in a fundamental way in the development of knowledge (1993, p. 1).

Phenomenology is the study of the structures of consciousness as experienced from a first-person point of view. One of its aims is to describe the structures of consciousness (such as perception) which allow acquisition of knowledge about the world, and to describe
the objects of experience as they are immediately present in one’s consciousness. A relevant problem concerns the comprehension of the structure of the temporal flow of consciousness, which allows for the perception of enduring objects as they are experienced by the subject.

In the case of music – which is composed by individual sounds that are heard as a continuous connected whole – the listener is able to find relationships among sound events occurring in the acoustical environment due to specific perceptual mechanisms, cognitive principles and neural processes. Empirical research related to these processes may contribute greatly to the phenomenological investigation of the structures involved in music perception and the way subjects perceive enduring objects and constitute their musical experience in a subjective manner. In particular, this combination of perspectives in an analysis and discussion of the problem previously highlighted through the case of the man in the metro may contribute to a better understanding of the way human beings know the world in which they live, become familiar with it and learn how to react to the events that they encounter during experiences in the world. I will develop various arguments based on a combination of empirical studies and phenomenological perspectives directed towards better understanding the problem of the man in the metro through the next seven chapters. Finally, I will consider their philosophical implications.

The question of how the man is able to immediately perceive music is related to the study of musical comprehension in the acoustical environment and what understanding sounds as constituting music means. The simple fact that the man classifies what he hears as ‘music’ indicates a kind of comprehension of the auditory environment that presupposes a specific organization of the acoustic information. In the first chapter (Basic musical understanding), I introduce the problem of musical comprehension of the acoustical environment and consider the factors that may influence a listener’s perception of a particular sound experience as music. The basic idea is that listeners understand sounds as music when they organize them in particular ways on the basis of specific mental processes and perceptual abilities, one of which is expectation. Such organizational processes allow the listener to have ‘basic musical understanding’: a concept that I introduce to describe the
immediate perceptual awareness of a musical event occurring in a complex auditory environment.

I do not consider other complex levels of comprehension that involve specific knowledge, experiences, mental processes, concepts and beliefs (like the intellectual or aesthetic understanding of music). I focus on the immediate perceptual awareness of music, which is the condition for every other level of comprehension, and try to highlight some of the aspects that characterize this perceptual comprehension: the organizational character of perception; the perceiver-environment relationship; the influence of learning as well as the possibility of innate perceptual structures and abilities; the subjective aspect of musical listening; the influence of concepts and beliefs and the intentional movement that characterizes every act of consciousness.

In the second chapter (The constitution of subjective musical experiences) I take into account the problem of meaning in music and discuss the idea of music comprehension as a process of constitution of musical meanings. Every perceptual awareness of a particular sound experience (musical or otherwise) has a specific meaning for the subject. From a phenomenological point of view, every experience is already meaningful, and every time an object is present to one’s consciousness with a particular organization, it has a specific meaning.

By organizing sounds in specific ways, the man in the metro constitutes his auditory experience as having a particular meaning: firstly, that of being a musical experience. Secondly, his musical knowledge and past experiences determine a more complex meaning which that musical experience can have in that particular moment and context. This may be something familiar, emotionally powerful, or having a specific musical meaning (like being in sonata form or the song of a famous Canadian rock band). What is interesting at the perceptual level is that the man does not need to consciously reflect on his experience in order to constitute these kinds of musical meanings.

I do not discuss the specific type of meaning music expresses, but propose to consider the basic perceptual understanding of music as the process through which a subject constitutes her diverse auditory experiences as having particular meanings. It is
because of this basic musical understanding, for example, that the man in the metro is able to have an acoustical experience that acquires the meaning of ‘music’.

At the end of the second chapter, I introduce the expectation process as a mental mechanism which is highly influential in the process of constitution of meaning: an argument that I develop further in the third chapter (*The expectation process*).

The process of expectation is a specific cognitive mechanism which, based on research results and various theories about the perceptual organization of sound, seems to be very important for the structuring of sounds into musical forms. I argue that expectation influences the subjective meaning acquired by the man during his musical experience in the metro station. When listening to music, or in the case of any auditory event occurring in the environment, listeners experience a particular state of expectation that guides them to unconsciously anticipate future sound events. In its broader sense, expectation may be considered as a basic strategy of the human mind that reflects a *tendency* toward the future and which is based on previous experiences. In music, expectation has often been related to the affective and emotional reactions a listener experiences when listening to music.

Through an analysis of the most relevant theories and empirical studies about musical expectations, I try to highlight the main aspects of this process. The aim is to understand and describe its influence on the constitution of the man’s musical experience in the metro station, particularly in the case where he encounters an unfamiliar musical event. For this reason, different scenarios related to music listening are introduced and discussed. These include: experiencing music unexpectedly, listening to familiar music and listening to unfamiliar music.

In order to better understand the way listeners understand and react to different musical events, in the fourth chapter (*Conceptual distinction: expectation versus anticipation*) I introduce a conceptual distinction between expectation and anticipation. The idea is that by defining the differences between two terms which are often used synonymously in much of the literature related to music perception, some new insights may be acquired regarding the comprehension of the subjective musical experience. This distinction may also be helpful in clarifying what expectations are, as well as their role in
the musical comprehension of an auditory environment. In particular, it may better explain
the different levels of comprehension a listener might have during the perception of
familiar and unfamiliar music.

In a broader sense, expectation refers to a complex mental process which is
fundamental for the comprehension of music. In its narrow sense, it is a mental state of
suspense about what is going on, during which a range of probable events are expected to
happen. Anticipation may be seen as a particular moment during the process of expectation:
it is the moment in which the listener predicts in her mind how the music will go on. The
ability to form anticipation may be what makes a difference in her understanding or
misunderstanding of unfamiliar music. For example, if the man in the metro is not able to
create anticipations because the music he hears is unfamiliar or completely unexpected in
that context, it may result in a feeling of misunderstanding.

Before developing this distinction, I take into account another aspect of music
perception. Looking at the cognitive science of music, the process of constitution can be
defined as an organization of sounds into structural forms. How can we describe the
structure that listeners perceive while hearing music?

Some empirical research tries to explain the process by assuming the idea of
‘correct musical understanding’, which presupposes that the perceived structure
corresponds to the one that has been notated by the composer in the score or established by
the performer. This is part of a common view belonging to the Western Classical tonal
tradition for which a piece of music finds its identity in the score and listeners understand
the music only when they grasp the structural features as notated.

Such presupposition limits the notion of a subjective musical experience because in
the attempt to give a causal account, it fails to describe the real perceptual experience of
music and does not take into account those kinds of music which lack a precise notation.
Moreover, it does not consider the different factors, contexts and situations that influence
one’s way of hearing music.

For this reason, in the fifth chapter (The perceived musical structure versus the
notated/established musical structure) I try to make a distinction between perceived and
notated/established structure, showing that in order to constitute musical experiences at the perceptual level and to have a kind of musical comprehension, it is not necessary to organize sounds according to the structure which has been established by the composer or the performer(s). To accomplish this task, I take into account the notion of structural hearing and the ways in which this type of hearing is distinct from the idea of a ‘natural’ tonal hearing. I discuss the main theories which assume that listeners have to perceive the notated/established musical structure in order to understand music, the notion of ‘correct’ understanding of musical structure and some empirical examples of structural hearing.

The main idea I emphasize in this chapter is that listeners perceive music in a structural way, but in order to have meaningful musical experiences and understand sound events in musical forms (as the man in the metro), they do not necessarily have to grasp the notated/established musical structure. Their experience is something subjective and the structuring of music is highly determined by past experiences and knowledge. This opens up the possibility of a kind of immediate comprehension of unfamiliar music. From this perspective the process of expectation plays a very important role.

An aspect of the perceptual experience of music which is at the basis of the constitution of musical meanings is its structural unfolding in time. Listeners constitute their perceptual comprehension of music during its temporal development; that is, the moment-to-moment unfolding of sound events (a note, a noise or a sequence of sounds) that occurs in their living auditory environment. The phenomenological account of the temporal structure of human consciousness is helpful in order to better understand what the man in the metro mentally does to hear music and to be aware of his perceptual experience. In particular, this temporal structure is fundamental for the constitution of subjective musical experiences made by a listener in the present moment of his perceptual experience in the living world.

The temporal structure seems to develop at two levels of complexity. At the first level, due to the movement of retention-primal impression-protention (which will be defined and discussed in the sixth chapter), the listener is able to perceive a sequence of sounds as a meaningful unity. At the second level, memories of past experiences and
expectations about the future are involved in order to constitute the meaning of a particular experience.

In the sixth chapter (*Listening in the moment and the role of anticipation versus expectation*) I briefly discuss Edmund Husserl’s account of the temporal structure, and I propose the idea of ‘listening in the moment’ – the ability to follow and be involved with music in the present act of listening to it and during the temporal unfolding of the musical event – to describe the way in which the man in the metro constitutes musical experiences in relation to unexpected and unfamiliar sound events. Within this framework, I try to clarify the differences between protention, expectation and anticipation. With reference to the hypothesis I introduced in the third chapter – that is, that there is an important conceptual distinction between expectation and anticipation – I consider the ways in which expectation and anticipation, as defined, influence ‘listening in the moment’, the basic perceptual understanding of music and the constitution of the man’s musical experience in the metro.

Expectation and anticipation are both acts of consciousness, but they do not correspond to the perception of a future event because such an event is not yet sensorily experienced. We could define expectation as a mental state, corresponding to the *pre-presentation* of a future event which is not well defined. There is a kind of reference to this future event, but it belongs to a range of probabilities that are part of the indefinite possibilities of the temporal horizon and, therefore, not exactly known. Anticipation, on the other hand, indicates a specific event in the future. It may thus be understood as a ‘quasi-perception’ of that future event. It is a kind of mental representation resulting from a strong expectation and is based on a past experience with similar musical situations.

Expectation and anticipation are related to different bodily responses and cognitive states. For this reason, the man’s experience, the meaning it acquires and the responses to the music that he perceives in the metro are influenced in different ways by the expectations and anticipations (if any) that he creates in the moment. By discussing how these mechanisms influence the basic perceptual understanding of music and the subjective constitution of musical experiences, the aim is to support the hypothesis that listeners can
have a kind of perceptual understanding of unfamiliar music during the present unfolding of a musical event.

In the last chapter (*The case of timbre expectation*) I propose the perception of timbre as an example of ‘listening in the moment’ and further explain the role of both expectation and anticipation in this process with relation to timbre. Particular attention is given to the case in which the man in the metro hears unfamiliar music: timbre becomes an important cue for the organization of an unfamiliar auditory environment. As has been demonstrated by recent empirical works, the musical dimension of timbre is very important to the organization of sounds and their constitution into musical forms – particularly for non-musician listeners.

With reference to some contemporary music and music belonging to cultures that are very different from the Western tonal tradition, I analyse some perceptual mechanisms that are used during musical listening and influence not only the organization of timbre, but also the way sound events are understood as music: the perception of similarities between sounds, the phenomenon of sound localization, and the recognition of sound source. These three important factors allow for the creation of expectations and anticipation based on timbre, and a musical comprehension of unfamiliar and unexpected sound events in the present moment of the perceptual act of listening.
BASIC MUSICAL UNDERSTANDING

Introduction

In the following discussion I highlight a specific problem: how the man in the metro station is able to immediately perceive music in such a complex auditory environment. In phenomenological terms, this problem considers how the man is perceptually aware of a temporally extended object among the complex auditory environment of the metro. Such a problem is addressed by relating the musical comprehension of an acoustic event occurring in a complex auditory environment and what understanding sounds as constituting music means: comprehension is the condition under which a listener has musical experiences.

In this chapter I focus on the musical comprehension of auditory events that is developed by a listener at the perceptual level, and I introduce the concept of ‘basic musical understanding’. In its general sense, musical understanding is conceived as a process of abstraction of meaningful structures from a complex auditory and acoustical environment. It is based on one’s own knowledge and abilities. The problem concerns how such a process is realized at the perceptual level and how the cognitive structures involved influence the way a perceiver understands the auditory environment.

Focusing on musical listening, we can see that there are different kinds and levels of comprehension. As Malcolm Budd observes, people hear individual notes that make up distinct melodies, harmonies, rhythms, sections, and so on, and they also hear the interaction between these elements. This indicates a musical understanding that comes in degrees along a number of dimensions and is highly influenced by subjective experience. For example, your understanding of a given piece or style may be deeper than mine, while the reverse is true for another piece or style. I may hear more in a particular piece than you do, but my understanding of it may be inaccurate. My general musical understanding may
be narrow, in the sense that I only understand one kind of music, while you understand many different kinds (Budd, 1985b).

With the present analysis, I propose a study of the comprehension listeners have at the perceptual level, which is the primary condition for more complex understandings, and do not imply the necessity for the listener to have explicit theoretical knowledge in order to listen to and appreciate music; to have musical experiences and to listen to music with understanding, listeners do not need to possess explicit concepts or use technical terms such as ‘melody’, ‘dominant seventh’, ‘sonata form’, and so on. This terminology is used by theorists to describe musical experience and perception in a systematic and consistent manner but does not proceed listeners’ musical understanding. As Stephen Davis observes, “following the music with understanding is rooted primarily in the way we experience it as we hear it. That understanding is cognitively founded, as I have suggested above. But this is not to say that one verbalizes that understanding as one listens, supplying a commentary that runs alongside the listening experience it describes. Cognitively informed perception does not have to be accompanied by linguistic thought” (2007, p. 14).

I do not address other levels of comprehension, such as intellectual and aesthetic understanding, which refer to musical experiences that involve specific knowledge, concepts and different levels of consciousness. I focus on the kind of understanding that allows people to immediately perceive particular sound events occurring in a complex auditory environment as music. I do not take into account that level of comprehension that involves listening to music as part of, for example, a defined musical tradition which has particular aesthetic properties or technical features. Such understanding entails more complex meanings and concepts such as those concerning the architectonic and stylistic aspects of a musical composition that are inferred after conscious reflection or aesthetic judgements. More precisely, I focus on the immediate comprehension a listener has of an auditory event; a comprehension which does not require the mediation of conscious thought.

The comprehension a listener has at the perceptual level is fundamental in order to have immediate awareness of a subjective auditory experience carrying specific meaning.
Because of its immediateness, it does not necessarily involve the mediation of conscious reflection. It is this first and basic level of music comprehension that allows an immediate perceptual representation of music. With reference to the case of the man in the metro, the problem to be discussed is how a sequence of sounds comes to be perceived as music. In other words, how the man identifies, and therefore, understands certain sounds as constituting music.

Firstly, I discuss what it means to understand sounds as music, defining it as a particular type of auditory experience. This involves the problem of what music is, which depends on our comprehension of music, since the concepts we have about music influence the way we listen to and understand it at the perceptual level. In the context and for the purposes of the present work, I do not discuss ontological problems about music. My attention is focused on the way listeners organize and understand sounds as music at the perceptual level, and consider the key elements and processes that determine such experience. The basic idea is that listeners understand sounds as music when they organize them in particular ways and on the basis of specific processes and abilities.

Focusing on that musical understanding which does not require the mediation of conscious thought, comprehension may be considered as a mental act in which a number of elements are grasped together. Listeners have perceptual comprehension when they organize the information they receive from their interaction with the environment in specific ways, forming complex musical forms. We will see how this organization is characterized and the way it is realized at the perceptual level through the help of some empirical research in the field of cognitive psychology of music. At the basic level, such comprehension is determined by the ability to organize sounds according to subjective knowledge, experiences and cognitive abilities which are often unconscious. This is the first step towards a musical understanding that could give rise to subjective representations of music and determine different kinds of musical experiences.

It is important to note that listeners perceive musical events as happening in certain ways which are shaped by cultural knowledge, constrained by their cognitive and neural structures, and influenced by context and the listener’s intentions. For example, there is a
kind of comprehension when a listener goes to a concert and another kind when she does
not expect to listen to music in a daily context or situation, as the case of the man in the
metro.

Through the analysis of both theoretical works and empirical studies, I try to
highlight what the main features that characterize this process of perceptual comprehension
are in order to comprehend what the man in the aforementioned example mentally does to
perceive music in the metro and to define the concept of ‘basic musical understanding’. In
particular, I focus on the organizational character of perception; the perceiver-environment
relationship; the influence of learning as well as the possibility of innate perceptual
structures and abilities on musical comprehension; the subjective aspect of musical
listening; the influence of concepts and beliefs and the intentional movement that
characterize every act of consciousness.

Levinson’s notion of basic musical understanding.

The fact that the man in the metro classifies what he hears as ‘music’ which is
distinct from other auditory experience indicates a kind of comprehension of the acoustical
environment that presupposes a specific organization of the auditory information. I
introduce the idea of ‘basic musical understanding’ to describe the immediate perceptual
awareness a listener has of a musical event occurring in a complex auditory environment.

The problem of musical comprehension has been discussed extensively in
philosophical literature.3 An interesting theory is that of Jerrold Levinson. In opposition to
those philosophers who follow a formalist and architectonicist point of view, he proposes
an enhanced version of Gurney’s concatenationism.4 One of the issues at the basis of his

---

3 See, for example, Davies (1998; 2007); Kivy (2001); Levinson (1997); Scruton (1983; 1997).
4 From a formalist point of view, listeners understand music when they grasp its compositional and
architectonic structure. It presupposes a kind of ‘musical analysis’ while listening to a musical composition.
This view is supported by Davies (1998), Scruton (1997), Kivy (2001; 2002) and Hanslick (1857) among
others.
Edmund Gurney was an English psychologist and musician. His major work is The Power of Sound (1980), in
which he presented an interesting conception of the music listening experience itself. Levinson refers to
theory is: how do we listen to music? The answer is to be found in the “phenomenon of following music, that is to say, of attending closely to, and getting involved in, its specific movement, flow, or progression, moment by moment. That is to say, it is not so much a matter of thinking articulately about the music as it passes, or contemplating it in its architectural aspect, as it is a matter of reacting to and interacting with and the musical stream, perceptually and somatically, on a non-analytical, pre-reflective level” (Levinson, 2006, p. 505).

In his book *Music in the Moment*, Levinson (1997) refutes the notion (often implicit in the writings of many music commentators and theorists) that people understand music after they have acquired conscious insights about its large-scale form. He tries to explain the aural comprehension of extended pieces of music by appeal to a tacit, unconscious correlation between present passages or bits with earlier ones, rather than an explicit, conscious grasp of relationships of the broad-span sort. With his theory he argues for the idea that music is essentially perceived as a chain of overlapping and mutually involved parts of small extent, rather than either a seamless totality or an architectural arrangement. From this perspective, “Musical understanding centrally involves neither aural grasp of a large span of music as a whole, nor intellectual grasp of large-scale connections between parts; understanding music is centrally a matter of apprehending individual bits of music and immediate progressions from bit to bit” (1997, p. 13).

Focusing on the musical listening itself, Levinson starts his book by observing that music consists of a series of successive events, which cannot be apprehended simultaneously in a single perceptual act. During the perceptual experience, listeners cannot perceive the form of a musical composition as a whole, they can only conceive it (or perhaps imagine it, in a nonperceptual way) as a whole. It is not possible to have a single perceptual musical experience in its entirety. With reference to the listeners’ understanding of ordinary tonal music, what is crucial is the local movement from note to note and phrase to phrase, and the involvement in the musical progression from point to point. This does not

Gurney’s perspective as concatenationism, which is understood as an account of what is central and distinctive in the apprehension of music.  
For an introduction to Gurney’s theory see Levinson (1997).

See Levinson, 1997, chap. one.
mean that Levinson intends to render irrelevant anything concerning the piece as a whole. With familiar melody, for example, one in some sense ‘hears’ the whole at any point. “Hearing musical movement is necessarily hearing a sonic entity not all of which is sounding at any instant, while at any instant, one hears the sounding notes as belonging to a musical flow, or as contained within a musical process, of which they form a part” (Levinson, 1997, p. 15). Thus, according to Levinson musical movement can be considered as a whole developing in time, which listeners cannot perceive in a single act. To explain this idea, he develops Gurney’s perspective on the problem and proposes the concept of quasi-hearing, which is “a sort of vivid apprehension of a musical unit, which goes beyond what is strictly heard, but stops well short of merely intellectual contemplation of a recollected event” (1997, p. 15). From this point of view, the architectonic contemplation of the musical structure is a different kind of musical understanding than that which is implied in the quasi-hearing of its moment-by-moment progression. In the latter case, listeners temporally maintain sounds which have just occurred and predict that which is about to occur without conscious efforts to do so. In this way, they perceive a quasi-hearing section as belonging to the same musical phenomenon.

According to Levinson’s perspective, listeners perceive music as a developing process and the sense of progression and continuation felt at each point is fundamental. This perceived movement plays a large part in the comprehension of music at the basic perceptual level. Levinson observes that “the listener is focused on individual parts as they occur, and on the connections of such parts with immediately preceding and succeeding parts. If appropriate expectations and response capacities are in place, focus of this local sort seems to suffice for comprehending listening” (1997, p. 25-26). The relationships created between sounds are therefore fundamental in order to understand music.

One could object to the notion that music comprehension is not an atomistic affair, as it seems from this perspective. But it is clear when he asserts that the earlier portions of a piece, through being heard, influence apprehension of later parts not contiguous with them that this is not what Levinson wants to say. This means that all the parts are related into a

---

6 Gurney emphasizes that the experience of envisaging a piece as a whole, when it has been retained in memory, is really nothing like the experience of music as heard, as registered and responded to in real time.
meaningful unity through an unconscious recollection. Moreover, concatenationism does not deny that the organization of a piece as a whole has effects on its parts. Large-scale reflection can facilitate the achievement of basic musical understanding, but it is not a necessary condition. The problem concerns what role deliberate attention to large-scale form plays in generating understanding of music as heard.\textsuperscript{7} Levinson asserts that the large-scale organization of a piece never becomes an object of perception itself, even if it affects the listener’s experience in some manner. In Levinson’s view, in order to basically understand a piece of tonal instrumental music, “it is unnecessary to possess any explicit knowledge of the piece’s formal structure or to maintain any awareness of its large-scale form. The only thing absolutely requisite is listening itself, grounded in sufficient prior exposure to related music and iterated sufficiently for the specific texture and movement of the given piece to be followed and responded to in its full individuality” (1997, p. 173).

Music lives and dies in the moment; it comes alive in the notes themselves, and musical understanding only requires exposure. Thus, as noted by McAdams and Bigand (1993, p. V), in order to hear music and understand an acoustic experience as musically meaningful, “all that seems to be needed is simply to ‘open one’s ears’”.

In this way, Levinson objects the supposition that listeners must acquire formal or structural knowledge (about large-scale structural relationship) in an explicit manner (architectonic awareness) to achieve musical comprehension, even of extended compositions. His point of view addresses the unconscious aspect of music comprehension and the importance of differentiating between possessing knowledge of musical structure in an articulate manner and perceiving it experientially from one moment to the next. He considers the extent to which understanding may be embodied in the ability to respond, discriminate or reproduce music rather than the ability to categorize, label or describe it.\textsuperscript{8} Thus, Levinson proposes the idea of ‘basic musical understanding’, which can be attained

\textsuperscript{7} To reply to this problem Levinson considers, for example, the case of the sonata form (1997, p. 87): it seems that the awareness of being hearing a sonata form influences musical understanding. Large-scale reflections might enable one to listen in a certain way and focus on specific aspects of the musical composition.

\textsuperscript{8} As Levinson observes, “the structuring of the piece, absorbed and retained unconsciously by the attentive listener, itself suffices for the aural rightness and satisfaction in question, and does not seem susceptible to significant enhancement through contemplation of such formal structure” (1997, p. 105).
without any awareness whatsoever of overall structure. It reflects an aural way of grasping music, which is different from grasping its formal structure. It is a perceptual understanding for which the listener does not necessarily need to know the architectonic structure of the music being heard.

Levinson notes that: “understanding music, as a number of philosophers have recently observed, is fundamentally a matter of hearing it a certain way” (1997, p. 29). In his view, quasi-hearing is a way of listening to music which involves aurally connecting tones currently sounding, those just sounded and those about to come, and synthesizing them into a flow as far as possible at every point. This reflects a basic musical understanding, which is identified by a locally synthesized rather than globally synoptic manner of hearing. For this kind of understanding, it is important to follow the development of events in real time. Such a process is realized due to the temporal structure of human consciousness; a point which I will develop in the sixth chapter.

Understanding music thus reflects an ability to hear sound environments in a particular way, for which learning plays an important role. In particular, Levinson relates quasi-hearing to a knowing how.9 “To have basic understanding of a piece of music is to have the ability to hear it in a certain way; thus, somewhat liberally, it is to know how to hear the music, how to experience it in listening. Such know-how, if we may call it that, is built up from prior hearings of the piece in question, while resting on a background of hearing competence derived form the hearing of music in the given tradition generally” (Levinson, 1997, p. 30). It follows that basic musical understanding is based on a particular way of hearing or being engaged by music, and it is developed at a primary and experiential level. More precisely, it refers to the immediate awareness of music, which is based on the ability to follow and be involved by the musical movement. ‘Basic’ means that such understanding is essential to any apprehension of music; it is fundamental to any further musical understanding and central to worthwhile musical experience of any kind. In the case of Levinson’s work, the object of comprehension is the piece itself, which is an aurally experienced sequence of tones. With the case of the man in the metro, I do not focus on

9 Levinson also distinguishes the ‘knowing how’ from a “knowing that”, which refers to propositional knowledge.
listening to a piece of music in its entirety, but on an unexpected experience that could involve listening to very short parts of a musical piece. The man in the metro, for example, has a basic musical understanding in the sense that he is able to perceive and be immediately aware of a sequence of sounds occurring in the acoustical environment as music.

I agree with Levinson’s ideas considering the experiential aspect of musical listening and the influence of unconscious learning: listeners learn to aurally organize music according to the most recurrent structures of the music to which they are exposed. His theory also reflects an important aspect of musical understanding: it is a process which is built in time, during the temporal unfolding of a perceived musical event and thanks to the temporal structure of human consciousness (a topic which will be widely discussed in the sixth chapter).

I disagree with the idea that listeners need some specific competences which are acquired during exposure to a musical culture in order to understand music adequately. More precisely, I refute the presupposition that in order to understand music at both the perceptual level and unconsciously, listeners have to grasp the ‘correct’ musical structure.\textsuperscript{10} This is based on concepts such as ‘expert’ listener and ‘correct’ musical understanding. From this point of view, it seems that listening to music with understanding presupposes musical knowledge about its structure. In this sense, Levinson’s theory is close to a formalist point of view, which is an area I will elaborate on in the fifth chapter.

Unlike Levinson, I argue for a basic musical understanding which is highly based on a subjective way of hearing music. It is not necessary that the listener unconsciously grasp the ‘correct’ global structure in order to have perceptual comprehension of music and to be able to follow it in its temporal unfolding, as Levinson’s expert listener seems to do. However, an important aspect of ‘basic musical understanding’ is the sense of unity that it allows, which is based on the ability to perceive a sequence of sounds as belonging to a meaningful unit. In phenomenological terms, it corresponds to the perception of sameness

\textsuperscript{10} As we will see in detail in the fifth chapter, the ‘correct’ musical structure corresponds to the organization that has been imposed on a musical composition by the composer. It is distinct from the ‘perceived musical structure’, which can be identified with the unconscious organization of sounds that is at the basis of every awareness of music at perceptual level.
of an enduring object. During this movement that, as I will argue in the sixth chapter, is made possible by the basic and primary level of the temporal structure of human consciousness, different cognitive principles are activated in order to meaningfully organize sounds in time.

As a lot of empirical studies have shown, an important cognitive principle which allows the apprehension of momentary parts as extended units is grouping. During music processing, sounds are grouped together into small perceptual units within a single larger perceptual unit and thus are situated in relation to each other. In this way sounds are organized in time. In the coming paragraphs, I will discuss this and other relevant aspects of ‘basic musical understanding’ in order to propose a definition that may be helpful for the comprehension of the man’s musical experience in the metro.

The organizational character of music perception.

One condition that allows a listener to have basic musical understanding of an auditory experience is the organizational character of perception, which allows the organization of the acoustical environment into musical forms and structures. That is, it creates relationships between sounds, which results in musical meaning. A lot of theories and empirical studies relate the process of organization to the structural aspect of music. It follows that music comprehension depends on the organization of sounds in musical structures.

According to Lerdahl and Jackendoff’s book *A Generative Theory of Tonal Music* (1983a), we may relate the comprehension of music to two processes: the recognition of a piece of music as belonging to a specific culture or to a genre within a musical culture, and the ability to grasp the structural and hierarchical organization of sounds. In the first case, musical comprehension is the result of a complex process of organization which is related to a categorization that depends on different factors including cultural influences, past

---

11 This book proposes one of the most relevant theoretical models about the cognitive principles involved in music listening.
experiences, musical training, concepts, beliefs and intentions. This organizational process allows the listener to understand a piece of music within her own culture. For example, when a listener categorizes a piece of music as ‘jazz’, she understands it as having a specific musical meaning: it is ‘jazz’ music. Such comprehension is based on her knowledge and exposure to the music of her culture. For someone else the same piece of music may belong to another musical genre or be non-categorizable because, for example, it is a totally unfamiliar style for the subject who experiences it. The categorization of a musical event as ‘a piece of jazz music’ corresponds to different kinds and levels of comprehension. Here I do not consider all the ways through which listeners come to make this identification. I take into account the immediate perceptual awareness a listener has of a piece of music as ‘jazz’ (i.e. without conscious reflection), and the structural and hierarchical organization of sounds that allows this kind of comprehension.

In their theory, Lerdahl and Jackendoff try to explain how a listener understands the music of her own culture by providing an account of her musical intuitions. In their opinion, the central task of such a theory should be to explicate the mental organization a listener imposes on or infers from the physical signal. By ‘musical intuitions’ they mean the unconscious knowledge through which the listener organizes what she hears, which goes beyond simply registering surface features such as pitch, attack, duration, volume and timbre.

As I have mentioned above, an important cognitive principle that allows organization is ‘grouping’, which refers to the way sequences of sounds are grouped together according to their acoustical and temporal features. A limit of Lerdahl and Jackendoff’s book is that they did not develop empirical studies to support their generative

---

12 Because categorization among a musical genre is not always the same for all listeners, we may consider it as a subjective process which is dependent on knowledge and experiences. For example, we are able to categorize a piece of music as belonging to a particular style or musical idiom if we have had sufficient exposure to it. If I have never had experience with free jazz, for example, I am not able to place it under its correct category (it could also have different categorization among different cultures), but I still understand it as music and as having a specific meaning.

13 In their view “a piece of music is a mentally constructed entity, of which scores and performances are partial representations by which the piece is transmitted” (Lerdahl & Jackendoff, 1983a, p. 2). Here I do not consider the Generative Theory in detail – I refer to this in the next chapter -, but it is necessary to clarify that it is explicitly psychological because it does not concern the organization of music in and of itself, but the organization that the listener is capable of hearing.
theory. However, we can find a lot of papers in music psychology that have applied the principles of the theory to empirical studies (Deliège, 1987; Bigand, 1990b). Following one of the principal questions addressed by their theory – what sorts of musical organization do listeners hear? – the two authors have developed, in analytical form, an explicit musical grammar which is capable of expressing what the listener hears. In this way, they try to explain which cognitive mechanisms are used by a listener when she listens to a piece of tonal music. In particular, the authors underline the fact that during her exposure to music and without any particular musical training, a listener unconsciously acquires some musical knowledge which she uses in order to understand the music she hears. More precisely, she acquires the basic rules of the musical system used in her own culture.

In this way, the comprehension of music seems to be related to the use of some cognitive organizational processes and the possession of musical knowledge. Both allow the listener to find musical relations between sounds and to unconsciously infer musical structures from the surface. From this perspective, understanding music entails the ability to organize sounds in a structural and hierarchical way, which is the result of cognitive principles (in Lerdahl and Jackendoff’s terms, musical intuitions) that reflect unconscious learning of the basic grammatical rules of the musical idiom to which a listener has been most frequently exposed. In this way a listener is able to grasp the underling fundamental structure of a musical piece and identify it as belonging or not belonging to a specific musical culture or genre. In Lerdahl and Jackendoff’s words, “such a listener is able to identify a previously unknown piece as an example of the idiom, to recognize elements of a piece as typical or anomalous, to identify a performer’s error as possibly producing an “ungrammatical”

---

14 This grammar, which takes the form of a system of rules, models the listener’s connection between the heard musical surface and the perceived organization or organizations she attributes to the music. The kind of organization a listener attributes to a given piece, then, is not arbitrary but it is highly constrained in terms of an explicit formal musical grammar.

15 Considering those aspects of the musical structure that are hierarchical in nature, the authors have identified four distinct hierarchical structures which are simultaneously imposed on a passage of music (the musical surface): grouping structure, metrical structure, time-span reduction and prolongational reduction. In proposing this analysis of the musical structures, which tries to specify a structural description for any tonal piece, Lerdahl and Jackendoff refer to Schenkerian analysis.
configuration, to recognize various kinds of structural repetitions and variations, and, generally, to comprehend a piece within the idiom” (1983a, p. 3).

Another interesting book which could be helpful in defining music comprehension as a way to organize sounds is *Conceptualizing Music. Cognitive Structure, Theory and Analysis*, by Lawrence Zbikowski (2002). His work provides an exploration of the process of musical understanding, which is described as a process through which sounds are transformed into structures that make it possible to grasp music. The purpose of the analysis presented in Zbikowski’s book is to show how listeners’ understanding of particular musical phenomena can be characterized in terms of specific cognitive processes and structures, thereby connecting concepts concerning musical understanding to research in cognitive science. According to Zbikowski, when we use the term ‘music comprehension’, we refer to cognitive processes that are at the basis of the organization of musical materials. Without the action of these processes we cannot understand music, nor can we have musical experiences.

Zbikowski analyses three cognitive processes that are particularly relevant for the organization of music: categorization,\footnote{Categorization reflects our ability to categorize incoming environmental stimuli. It is a basic cognitive process through which we structure our thought, and categories are basic structures through which we understand the world. For example, to recognize a book is to identify it as a member of the category book; to recognize a tree is to identify it as a member of the category tree. Categories are not given by nature and they are subject to change and modification. They are in fact determined by the ways in which humans interact with their environment. Categories shape our understanding of phenomena, as it happens in the case of consonant and dissonant intervals. As Zbikowski observes, “consonance and dissonance are not naturally occurring properties, but ways of constructing an understanding of musical organization” (2002, p. 155).} cross-domain mapping\footnote{Cross-domain mapping is a process through which we structure our understanding of one domain (which is typically unfamiliar or abstract) in terms of another (which is most often familiar and concrete).} and the use of conceptual models.\footnote{Conceptual models consist of concepts in specified relationships, which pertain to a specific domain of knowledge. They provide the first level of organization for concepts.} Explaining the role that these basic cognitive structures play in the understanding of music, he also shows how the mind works in organizing other types of knowledge about the world.

As in the case of Lerdahl and Jackendoff, the structural organization of musical elements represents a fundamental process for comprehension of the sound environment as
music. In their work *The Angel of Death*, McAdams, Reynolds and Battier (2005) relate the comprehension of music to the way sounds are structurally organized and they consider musical understanding as a process through which sounds are transformed into structures that make it possible to grasp music. As we will see in detail, it depends on different processes such as the perception of similarity between musical materials. Such perception underlies a large part of a listener’s musical experience, including associations made between themes or motifs and their variations, the formation of musical categories and the sense of familiarity.

From the point of view of music theorists and ethnomusicologists, the organizational process is fundamental for comprehension of music. As John Blacking observes, “musical execution, just because it is different from the production of noise, is inconceivable without the presence of an order between sounds” (1986, p. 33). To underline the relevance of organization at different levels of musical activity – composition, execution and listening – 20th Century composer Arnold Schönberg affirmed: “without organization music would be an amorphous mass, as unintelligible as an essay without punctuation, or as disconnected as a conversation which leaps purposelessly from one subject to another” (1967, p.1).

In order to better understand the relevance of the organizational character of perception, many researchers have conducted analyses of people suffering from brain damage. In relating sounds, listeners build hierarchies of relevance and, also at neural level, temporal events are coded hierarchically. Some studies about the perception of music for which participants have suffered brain damage show that, in those cases where they are not able to relate sounds, it becomes difficult for participants to perceive structural and meaningful musical forms while hearing a sequence of sounds. Moreover, without the perception of musical relationships, they cannot generate expectations. A similar problem holds for participants with memory disorders, and difficulties may arise when they try to create expectations about future musical events.

19 In the paper *The Neural Processing of Complex Sounds*, Griffiths (2003) offers an example of this structural and hierarchical way of working through the analysis of the brain structures that are involved in temporal analysis.

One example of a type of brain damage that affected the perception of music is in the case of Rachel, which Oliver Sacks describes in his book *Musicophilia: Tales of Music and the Brain* (2007, p.21). The patient was not able to categorize sounds in relation to one another; she was able to isolate sounds, but could not integrate different musical elements into a musical piece. In his book, Sacks observes that “there are many levels in the brain at which perceptions of music are integrated and many levels, therefore, at which integration may fail or be compromised” (2007, p. 123). Rachel lacked the ability to produce musical imagery and suffered from working memory dysfunction, both which impacted her continuity in musical listening and composition. A Sacks notes: “with the difficulty in imagery comes a difficulty in working memory, and this makes it impossible for her to retain what she has just composed” (2007, p. 126-127). This also prevents the temporal structure of human consciousness from accomplishing its function of relating sounds in time. We will see in the sixth chapter that this is a fundamental aspect of music comprehension.

A question that might be addressed at this point is: what are the cognitive processes that are at the basis of the organizational character of music perception? As Lerdahl and Jackendoff indicated in their *Generative Theory of Tonal Music*, some cognitive principles that are at the basis of music perception have been well explained by *Gestalt* psychology. For example, from its surface structure listeners infer the first level of the structural

---

**Gestalt** psychology refers to theories of visual perception developed by German psychologists in the 1920s. It emphasizes that people perceive objects as well-organized patterns rather than separate component parts, and theorists attempt to describe how individuals tend to organize visual elements into groups or **unified wholes** when certain principles are applied. According to this approach, when we open our eyes we do not see fractional particles in disorder. Instead, we notice larger areas with defined shapes and patterns. The ‘whole’ that we see is something that is more structured and cohesive than a group of separate particles. The focal point of *Gestalt* theory is the idea of ‘grouping’, which describes how people tend to interpret a visual field or problem in a certain way. The main factors that determine grouping are: proximity, similarity, closure and simplicity. These factors are called the laws of grouping (for example, law of proximity and law of good continuation), which describe why some elements appear to go together rather than remaining isolated and independent. The principle of similarity, for example, states that things which share visual characteristics such as shape, size, color, texture, value or orientation will be seen as belonging together. Similarity (on which I will focus in the fifth and seventh chapters) thus occurs when objects look similar to one another. People often perceive them as a group or pattern.
representation of a piece of music (grouping structure) by using and organizing sounds on the basis of preference rules which follow the Gestalt principles of similarity and proximity.

Experimental studies that support perspectives based on Gestalt psychology have been developed by various researchers.\textsuperscript{22} One such important work conducted by Irene Deliège (1987) considers the grouping behaviour of two categories of subjects (nonmusicians and musicians) in order to discuss the validity of the proximity rules and the change rules – both of which express the intuitive organization of groups in music perception in relation to the Gestalt principles.\textsuperscript{23} The work highlights an aspect of music perception which has important consequences for the comprehension of music: without the organization of sounds, a listener would not be able to perceive a sound environment as musically meaningful.

The man in the metro unconsciously analyses the acoustical environment through his auditory system, which includes both physiological and psychological mechanisms. He extracts some information and organizes it into structured units. According to both the Gestalt principles and the acoustic features of sounds, he groups sounds that are similar in pitch and timbre, and those which are spatiotemporally proximate. These processes reflect both an automatic tendency to organize sounds and the use of learned structures and culturally determined schema, thus resulting in musical organization and comprehension of sounds.

The perceiver–environment relationship: Clarke and the ecological approach to musical listening.

A theory that might be helpful in comprehending the man’s musical experience in the metro and the kind of understanding that he has at the perceptual level has been presented by Eric Clarke in his \textit{Ways of Listening} (2005). He tries to describe the way in

\begin{itemize}
  \item \textsuperscript{22} See, for example, Deliège (1987), Deutsch (1982a).
  \item \textsuperscript{23} For an introduction to the Gestalt theory see Wertheimer (1923); Köhler (1961); Ehrenfels (1988).
\end{itemize}
which musical sounds take place in a wider context of other sounds – as in the case of the man in the metro, who is exposed to a wide range of different sounds. Musical sounds inhabit the same world as other sounds, and people use the same physiological system during their exposure to the auditory environment. How are they able to distinguish and understand some sounds as musical or as constituting music?

According to Roger Scruton (1997), we can make a distinction between the ’acoustical experience of sounds’ and the ’musical experience of sounds’.

In this way he distinguishes between the sounds of the ‘everyday’ world and the tones of music, which seem to exist only in the mind of the perceiver. Clarke goes on to assert that “Scruton’s approach is a consequence of the fundamental claim that musical events are “secondary qualities” – not tied to, or a part of, the physical circumstances of the real world but separated from them, and capable of behaving in ways that are not constrained by the real world” (2005, p. 68-69).

Even if the physical and acoustic features of sound are very important, McAdams and Bigand observe that “our perception of the sound world thus greatly surpasses the quality of the sensory information available at each instant: it results from mental processing” (1993, p. 2). This seems to be a different perspective than the one which is taken by Clarke. Since the beginning of his book Clarke maintains that he disagrees with the information-processing approach to perception, for which the structure of music is not in the environment, but rather imposed on an unordered or highly complex world by the perceiver. Clarke takes an ecological approach to music perception and, making reference to the ecological perceptual theory proposed by James Gibson (e.g. 1966; 1979), his aim is “to discuss the ways in which listeners interact with the general auditory, and more specifically musical, environment: to discuss listening to music as the continuous awareness of meaning, by considering musical materials in relation to perceptual capacities” (Clarke, 2005, p. 5).

---

24 “The starting point for Scruton’s argument is the distinction between sound and tone, which he identifies in three specific attributes: the distinction between what he calls the “acoustical experience of sounds” and the “musical experience of tones”; the distinction between the “real causality” of sounds and the “virtual causality” that relates tones to one another in music; and finally the distinction between the sequence of sounds and the movement of tones that listeners hear” (Clarke, 2005, p. 68).

25 Ecological research concerns the study of organisms in relation to their environment.
The ecological approach emphasizes the structure of the environment itself and regards perception as the pick-up of that already structured perceptual information. According to Clarke, perceivers are highly structured organisms capable of adapting to the auditory environment. However, studying the way listeners understand the acoustical environment in musical forms may be helpful in explaining their perceptual awareness of meaning and how they ‘hear sounds as’ music.

With the information-processing approach Clarke agrees with the idea that musical perception develops as a set of stages or levels which range from psychoacoustics to more cognitive aspects to the abstract structures of larger stretches of music. In opposition to an architectonicist perspective, Clarke underlines that “listeners are more immediately aware of the simple features of musical sounds than they are of higher-level characteristics, which are regarded as more abstract, complex, and remote” (Clarke, 2005, p. 15). He also disagrees with the emphasis given to the idea of mental representation (both as the final state that the system achieves, and as intermediate stages along the way) and the disembodied and abstract aspect of the information-processing theory (as well as the score-based and abstract structuralist music theory).

In following Clarke’s hypothesis, it is necessary to consider what is directly specified by environmental information: “this information directly specifies properties of the object itself to an organism equipped with an appropriate perceptual system” (Clarke, 2005, p. 18). An interesting concept that he proposes is that of ‘resonance with the environment’: the listener needs a perceptual system that will resonate to the environmental information, adapting to its features and changes. The resonance of a perceptual system with its environment is a product of evolution and adaptation.

Clarke’s theory is characterized by three main factors:

1. the relationship between perception and action: resonance is not passive, but an active process in which the perceiver is constantly engaged with her environment.
2. adaptation: organisms and their environments are constantly changing and their relationship is the product of mutual adaptation.
3. perceptual learning: perceptual features are not fixed. Exposure to the environment shapes these perceptual capacities and enculturation determines comprehension (in a way that listeners learn to understand the environment).

At first sight it seems that the ecological approach denies any kind of interpretation of the environment: if perception is related to the capacity to grasp the structures of the environment, then the listener understands music only when she grasps an already defined musical structure. But what does it mean to ‘grasp the structure of music’? What is a musical structure? In relation to these problems, it seems that there is a contradiction in the ecological approach because on one side it claims that music (as part of the auditory environment) has its own structure but, on the other, it stresses the development of subjective ways of listening to music. He observes that “the ecological approach presents perception as a mutual relationship between organism and environment, so that every description of perception is therefore specific to an individual’s capacities and perspectives” (Clarke, 2005, p. 156).

Two main problems arise with Clarke’s theory: firstly, if the structure is in the environment, then this structure must be the same for all listeners. But how can individuals develop different ways of listening and have different musical understandings if this is the case? Secondly, if comprehension is influenced by experience and subjective knowledge (perceptual learning), does it follow that if a listener does not grasp the ‘correct’ structure of the environment she does not have comprehension at all? A related problem concerns the definition of ‘correct structure’ and what ‘correctly perceiving it’ means. In Clarke’s opinion, an example of correct perception is the perception of a chord as a single entity by untrained listeners. In my opinion, this seems to be a vague and simple way of discussing the problem and does not explain what musical understanding is. A study of music perception requires more distinction and clarification than Clarke has provided.

I argue that a listener has basic perceptual understanding of music even if she does not grasp the ‘correct structure’. It is a kind of comprehension which reflects a different

---

26 Music perception could also be understood as an interpretation of the sound environment: through its characteristic organizational processes it allows the listener to understand music, to grasp the shape of the experienced musical event and give rise to meaningful representations of music.
degree of music comprehension from that of other listeners as well as a different way of perceiving music. I will discuss in detail the problem of the perceived structure and its comprehension in the fifth chapter.

An interesting aspect of Clarke’s theory is the relationship between perceiver and environment, which could explain both why listeners hear the same musical event in different ways and how they are able to create a musical communication based on common knowledge and structures. Perception is, in Clarke’s view, a reciprocal relationship between perceivers and their environment. A central principle of the ecological approach concerns the notion that “perception must be understood as a relationship between environmentally available information and the capacities, sensitivities, and interests of a perceiver” (Clarke, 2005, p. 91). As we have seen, according to Clarke the sound environment has its own structure and the perceiver has the adapted perceptual abilities to internalize this structure. But this does not necessarily imply that in cases where listeners do not grasp the structure of a piece of music they are not able to have any kind of comprehension or meaningful experiences. In the interaction between perceiver and environment there is always a kind of organization that is based on both acoustical features and perceptual abilities. Such perceptual abilities are developed by individuals in different ways because of the influence of different cultural and experiential factors such as, for example, culture, training, age and context. In order to better understand this point, it could be helpful to discuss how such a relationship is developed and which perceptual processes are involved.

The first important aspect of the perceiver-environment relationship is the listener’s point of view in the world. Considering the relationship between music and listener, Clarke uses the expression “subject-position” to indicate the way in which features of music shape the general character of a listener’s response or engagement with music. Even if each listener has personal experiences, background and attitudes that influence the way they listen to music, there is, in Clarke’s opinion, a limit to this potentially infinite plurality that is determined by the properties of music itself. In this way Clarke refuses both unconstrained relativism of reader-response theory and rigid structuralism. It follows that the acoustic and temporal features of the environment influence the way listeners organize
music, who, due to the particular organizational and compositional features they are accustomed to hearing in music within their own culture, develop specific ways of hearing music.

In order to understand the reciprocity that characterizes the perceiver-environment relationship and its influence on ‘basic musical understanding’, we must not focus only on environmental features, but also on the perceptual abilities that determine the subjective perceptual experience and the listeners’ predispositions and habits. The analysis of some cognitive studies in the forthcoming chapters has the goal of understanding both the way perceivers adapt their perceptual structures to environmental features and the influence of their ‘subject-position’, and how they develop subjective ways of listening the same musical piece. Thus, perception is understood as being based on a relationship between environmentally available information and the capacities, sensitivities and interests of a perceiver.

To underline the relevance of the perceiver-environment relationship in basic musical understanding, it could be helpful to take into account the concepts of ‘point of view’ and ‘being-in-the-world’, as developed by Merleau-Ponty (1942; 1945). Because she has a physical body, the human being is in constant and direct relation with the environment. She is not an external entity who imposes her cognitive and perceptual structures to the world outside her, but rather lives inside the world, constantly shaped by it and shaping it. She lives in a particular place in the world and she experiences and interacts with it from her subjective perspective. The musical experience arises from a relationship with the living world and depends on this subjective point of view in the world.

As Gibson observes, the relationship between perceiver and environment is neither a case of organisms imposing their needs on an indifferent environment, nor is it a fixed environment determining possibilities. Listeners have different responses in relation to their points of view and learned knowledge. This implies, for example, that they reply in different ways to the question ‘what do you hear?’. On the basis of the listener’s way of listening, the same sound may be most immediately heard as the sound of a cello; as Western art chamber music; as a G# rising to an A or as the opening of the first movement.
of Beethoven’s Op. 132 string quartet (Clarke, 2005, p. 158). This example demonstrates that there are many different ways of understanding a musical event.

The main features of the relationship between perceiver and environment that have been underlined by Clarke’s ecological approach are: the dynamic aspect of perception, which involves motion; the plasticity of the brain and the related openness to new information; the relevance of learning and the role of information. Concerning this first feature, perception is a dynamic process through which the perceiver is in constant communication with the environment and it involves a sense of motion. In music, Clarke writes, “the sense of motion or self-motion draws a listener into an engagement with the musical materials in a particularly dynamic manner” (2005, p. 89). This reflects an intentional movement toward the environment that characterizes all the human life. As we will see in the next chapter, this intentional movement contributes to the constitution of meaning. Addressing Clarke’s second point, the plasticity of the human brain is fundamental to an individual’s ability to react to the constant changes that occur in the environment. Finally, with regard to the third feature outlined by Clarke, learning has a central role in music perception and comprehension. Together with the flexibility of perception and the plasticity of the nervous system, learning allows for immediate adaptation and reaction to the changing environment. “Perceptual systems become attuned to the environment through continual exposure, both as the result of species adaptation on an evolutionary time scale, and as the consequence of perceptual learning within the lifetime of an individual” (Clarke, 2005, p. 25).

Cognitive empirical studies about the perception of tonality agree with Clarke’s hypothesis in that learning the regularities of the auditory environment is an important aspect of music perception and musical understanding. As Clarke observes, people become attuned to tonal music through simple exposure, due to the interaction of the regularities of the tonal environment with certain fundamental perceptual capacities of the auditory system. Listeners learn the regularities of the environment because their perceptual system is sensitive to environmental invariants. The information they learn is the result of a need to know ‘what is going on’. From this perspective, the environment is seen as a source of
information. In Clarke’s view, the perceiver has adapted her perceptual structure and abilities in order to pick-up the structures of the environment. He writes: “the ecological approach to perception emphasizes the perceiving organism’s adaptation to its environment, and the manner in which perceptual information specifies events in the world” (Clarke, 2005, p. 154).

In this way the listener develops a kind of understanding that depends on the way she interacts with the environment as her source of information. It follows that comprehension of musical structures in the environmental is a subjective experience because it depends on the perceiver’s learned knowledge and abilities, as well as the context and situation in which music is experienced. The way people listen to music during a concert, for example, is different from their everyday listening and unexpected encounters with musical events. In the former case they have a ‘musical listening’ (that we may call ‘structural listening’) that can be understood as the perception of musical event structures specified by sound. It involves the listener’s attention to the qualities and properties of musical sounds in themselves, and their purely sonorous relations with one another. In the latter case the perceiver’s way of hearing involves detecting objects and events in the world which are specified by sounds. In the case of the man in the metro, this occurs at the perceptual level in order that he may be immediately aware of the events that are occurring in the complex auditory environment of the metro. The two cases exemplify different ways of listening in relation to the context and situation that is experienced by the perceiver.

In Clarke’s opinion “perception always involves the reciprocal relationship between the opportunities of the environment and the capacities of the perceiver” (2005, p. 148). Musical understanding develops from this relationship. At the same time, we have to keep in mind that “the directness of our perception of the world is not an inexplicable or “magical” reciprocity between perceiver and environment: it is the consequence of adaptation, perceptual learning, and the interdependence of perception and action” (Clarke, 2005, p. 47).

At this point we may argue that the perceiver-environment relationship is fundamental to the listener in order that he or she develop basic musical understanding of a
sound event occurring in the auditory environment. The organizational process that allows such comprehension is highly influenced by the features of the environment in which the subject lives; to be in relation with the environment is therefore fundamental. However, this process of knowledge is not only determined by the structures of the acoustic environment, but also by the perceptual structures and abilities of the perceiver, as well as the limits that constrain such abilities. These are shaped by experience and they depend on the point of view of the listener. In this way, she develops a subjective manner of listening to music and understanding the auditory environment in which she lives.

To conclude, I follow the idea that music perception develops as a set of stages or levels. Listeners do not ‘impose’ their perceptual structures on the environment from an external position, but they use them in order to organize the information they receive in their constant interaction with the auditory environment.

The influence of learning in the perceptual organization of music and the possibility of innate perceptual structures and abilities.

An important factor that influences ‘basic musical understanding’ is the process of learning the regularities of the auditory environment to which one is exposed. This allows the listener to acquire knowledge, which is unconsciously used to analyse and immediately organize sensory data. In this way, she becomes familiar with the living world. As McAdams and Bigand observe,

Acquired knowledge interacts with the current sensory data to interpret the auditory stimulation. Imagine for an instant that we are being guided through an Amazonian rain forest: we would hear exactly the same noise as the native of the region that accompanies us but we would be incapable, due to a lack of knowledge of the environment, not only to extract from the sound background events corresponding to the cries of iguanas and Macaques, the songs of Wistiti monkeys, or the rustling of the leaves of tropical trees, but we would also be unable to assign meanings to the entire sound structure that may in the long run be important for survival. In the same way, the Golden Ears of the French Navy have been trained in sonar detection to listen to what most of us would hear as a noisy underwater sound field and to perceptually segregate and identify a multitude of underwater sources such as clicking shrimp, whales, porpoises, schools of fish, and ocean-going vessels. They even
succeed in classifying such vessels as commercial or military, as surface or submarine, as diesel or nuclear, as Russian, American, or French. In particularly dramatic cases, an inability to perceptually segregate sound sources and recognize their nature could fail to lead a listener to deduce from these signals the presence of imminent danger: the appearance of a jaguar, or a clandestin submarine, for example. (McAdams & Bigand, 1993, p. 2)

The perceptual awareness of auditory events as having specific meaning is therefore determined by the knowledge the perceiver acquires during her experience and simple exposure to the auditory environment (particularly in the case of the Amazonian rain forest, the case of the French Navy required conscious training). In musical listening, theorists such as Lerdahl and Jackendoff, Meyer and Huron, for example, have often related the process of learning to the cultural, syntactical and grammatical rules of the musical genres to which listeners are most exposed. Such knowledge, which could be consciously or unconsciously learned, allows them to organize sounds into structural musical forms.

In his paper *Vers la formalisation des processus impliqués dans la comprehension musicale*, Emmanuel Bigand (1990b) offers his approach to the problem of music comprehension and the role of musical knowledge. He begins the analysis with an analogy between music and language.27

Referring to Chomsky, Schenker, Lerdahl and Jackendoff, he highlights that the common feature between music and language is the ability to create an infinite number of combinations with a finite number of rules.28 From this perspective, music is characterized by a system of rules used to create infinite musical possibilities.

Once listeners have learned this system of rules, they are able to organize sounds into musical forms and have meaningful musical experiences. They are also able to grasp the meaning a particular piece of music has within a culture. In this sense, music comprehension is a process of abstracting musical structure, which is accomplished by inferring information from the superficial aspects of music through the use of a learned musical syntax. In other works, Bigand underlines the fact that listeners use their musical

---

27 The analogy is based on the idea that: “la musique et le langage son deux formes de communication auditive qui impliquent l’existence d’une compétence spécifiquement humaine à communiquer par l’intermédiaire de formes sonores. Les activités dérivant de cette compétence doivent donc vraisemblablement contenir un certain nombre de points commun” (Bigand 1990b, p. 86).

28 Bigand makes reference to the idea of ‘transformational generative grammar’, which is supported by Chomsky (1970); Lerdahl & Jackendoff (1983a); Schenker (1935).
knowledge to organize sounds and understand what they are listening to.\textsuperscript{29} Thus, he highlights the influence of the culture to which a listener belongs to: culture determines the way a listener organizes sounds, and then determines her comprehension of music.

The first problem with Bigand’s perspective involves the explanation of what ‘musical knowledge’ is. He seems to assume that musical knowledge requires understanding of a system of grammatical rules concerning structural aspects of music which are highlighted by music analysis. The major question is: in what form does one have comprehension of these rules? Here I think it is necessary to distinguish between ‘formal’ musical knowledge, which is learned through conscious training, and ‘informal’ or ‘experiential’ knowledge that all listeners (both musicians and non musicians) unconsciously learn through mere exposure to the music of their culture. The effects of unconscious learning are evident in some experiments and studies that have been conducted with infants, as their behaviour in response to repetitive music listening shows an ability to learn structural aspects of music without formally understanding it.\textsuperscript{30} Lerdahl and Jackendoff also note the unconscious musical knowledge that expert listeners\textsuperscript{31} learn through exposure. Such knowledge allows the listener to organize sounds into musical forms and become sensitive to the musical structures of the style they are most exposed to. For example, she becomes able to recognize rhythmic changes that are due to structural variations or to be sensitive to structural errors that are due to an incorrect use of the grammatical rules (such as when the performer does not conclude a cadence with the correct chord).

Another problematic aspect of Bigand’s account of music comprehension is represented by those musical genres or styles that are not based on a defined system of rules. Bigand, as well as Lerdahl and Jackendoff and other authors, only refers to Western tonal music, which is generally highly structured by a system of grammatical rules. But

\textsuperscript{29} See Bigand 1993; 1990a.
\textsuperscript{30} See, for example, Threub (1987); Schellenberg & Threub (1996); Saffran, Aslin, Newport (1996); Saffran, Johnson, Richard, Newport (1999).
\textsuperscript{31} In their view, an ‘expert listener’ is a listener who is very familiar with tonal music thanks to simple exposure.
what about a genre such as free jazz improvisation, which not only refuses tonality, but also the use of strong and regulated structural relationships?

The process of learning does not necessarily imply that the musical knowledge one acquires during his musical experience coincides with the rules of a specific musical genre or those which a composer might have consciously employed when composing a piece of music. Additionally, the way music is perceptually organized does not necessarily correspond to the way the composer organized it in the score. I will return on this topic in the fifth chapter.

The point to underline here is that learning is a subjective process: it is determined by the regularities of the environment and influenced by shared musical conventions, but the knowledge one acquires in listening to music reflect her subjective experience. It follows that a listener organizes the same piece of music differently from other listeners because she has different musical knowledge and a different point of view in the world. For example, if a person who has always listened to classical tonal music attends for the first time a free jazz performance, she understands it differently from a person who is used to listening to that kind of music. Therefore, in the case of unfamiliar music for which the listeners do not know the grammar (if any), they organize it according to their knowledge. In this way, they develop a subjective musical listening.

As David Huron observes, when we listen to music that is different from what we are accustomed to in our own culture (for example Tibetan music) we feel that we somehow get the gist of it (Huron, 2006, p. 376). In this case, what does it mean to understand the music of another culture? Musical understanding can involve many perceptual processes and be influenced by different factors, and listeners can have different kinds of comprehension because of their learned knowledge. There is not only one ‘correct’ way to understand music; moreover, “the identification of the sources of sound can be different according to the kinds of information to which we pay most attention: regularities in the noise or the timbral qualities of the materials being set into vibration. Faced with ambiguous stimuli of this kind, the perceptual system normally makes a subconscious decision in order to organize the sound figure. The knowledge possessed by the listener
certainly plays an important role at this level” (McAdams & Bigand, 1993, p. 2). This viewpoint considers music to be alive, dynamic, always changeable and subject to different interpretations.

Although Huron maintains that individual knowledge plays an important role in musical processing, his ITPRA (Imagination, Tension, Prediction, Reaction and Appraisal) theory also shows that there are some common mechanisms at the basis of the psychology of expectation that are shared by people of different cultures:

the ITPRA theory assumes that all the world’s people share a common biological heritage, even though the specific forms of expectation may differ radically between cultures and over time. That is, the theory presumes that everyone exhibits the same types of memory, the same inductive learning processes, the same tendency to form multiple mental representations, the same physiological responses to surprise and habituation, and the same imagination, tension, prediction, reaction, and appraisal components. (Huron, 2006, p. 376)

Why, then, do we have different understandings of music? In order to come up with an answer we have to consider the structures of individual minds and try to better understand how the human mind works.

Some interesting research findings in the field of neuroscience support the idea that listeners subjectively learn knowledge that they use to organize sounds into musical forms.\(^{32}\) Studying the modular organization of music cognition and comparing the effects of procedural vs explicit music teaching on brain networks, Altenmüller and colleagues demonstrate that cortical activation during music processing reflects the auditory ‘learning biography’. This is an interesting concept that indicates the personal experiences accumulated over time by a listener. In listening to ‘real music’ (i.e., not constrained by laboratory conditions and experienced in real contexts), “we have to expect individually formed and quickly adaptive brain substrates, including widely distributed neuronal networks in both hemispheres” (Altenmüller, 2003, p. 346). This indicates not only subjective listening and learning at the neural level, but also an adaptive and dynamic type of listening that is realized through relations with the environment.

\(^{32}\) See, for example, Altenmüller (2003); Bever & Chiarello (1974).
To understand the influence that learning has on music comprehension in the case of unfamiliar music and to underline this as a subjective experience, an interesting point of view is again that of David Huron. As Huron observes, learning is not tied to a rigid system of rules, but it is related to the regularities of the environment which shape the perceiver’s knowledge. The organization of sound is built on a set of probabilities that reflect such knowledge and that are related to one’s own experience in the world. Learning is related to a statistical organization which is based on the probabilities that something which happened in the past will occur again. As I previously noted, not all musical styles make use of a precise musical grammar. Listeners in these cases refer to past experiences, repetition and probabilities in order to organize the music they hear.

In general, we could say that learned musical knowledge involves musical schema and structures that are acquired thanks to past experiences, exposure and repetition of particular musical patterns; some of which reflect grammatical rules because they occur often. Such a system of knowledge is flexible, dynamic and open to new situations. This is a fundamental aspect which allows, together with the perceptual structures and processes involved, listeners to organize incoming auditory information in a meaningful way even if the music is unfamiliar or lacking a known musical grammar.

This does not mean that music is exclusively a subjective experience. Through music people express and communicate something. Learning is important not only in order to understand music, but also because it results in the acquisition of musical regularities within a shared culture or style-specific musical context. Such acquisition results in the creation of a musical communication. This implies the possession of similar perceptual processes and structures in order to find those regularities. The inability to find such regularities because of, for example, brain damage, shows how important the possession of similar perceptual structures is in order to process sound information and create a musical dialogue.

To sum up, the listener’s comprehension of a sequence of sounds as music depends on her past musical experiences and the musical knowledge she has acquired during exposure to the musical style of her culture. Such a system of learned knowledge is not
fixed but rather continually altered by new experiences. It follows that basic perceptual understanding of musical events is something learned, not innate, and subjectively constituted through experience. Bertrand Russel claims: “Understanding music is not a matter of dictionary definitions, of knowing this, that, or the other rule of musical syntax and grammar, rather it is a matter of habits correctly acquired in one’s self and properly presumed in the particular work” (as cited in Meyer, 1956, p. 61). From this point of view, the perceiver-environment relationship that has been underlined by Clarke is fundamental in order to shape one’s own musical knowledge.

Some questions arise in relation to the way such a learning process is developed by the human mind and through which cognitive and neural processes occur. Musical knowledge is not innate, but I propose that some cognitive organizational principles and structures that allow the learning of such knowledge could be considered as common features of the human mind. Huron, for example, referred to memory, inductive learning processes, expectation and other processes as common biological mechanisms shared by all people. As Davies observes, “the capacity to recognize music as such is not purely biological or natural. But also, because we can often identify the music of foreign cultures as music, even if we are not well situated to appreciate and follow it because we have no grasp of the principles on which it proceeds, neither is the identification of music purely or arbitrarily cultural” (2007, p.5).

At the end of their book Lerdahl and Jackendoff, with reference to Chomsky (1970), propose the hypothesis of ‘musical universals’. These principles are available to all listeners and allow them to organize the musical surfaces they hear, no matter what idiom they are experienced in. For example, a universal aspect of musical listening is that musical intuitions are organized along four hierarchical dimensions which Lehrdal and Jackendoff outline in their book. They also claim that “much of the complexity of musical intuition is not learned, but is given by the inherent organization of the mind, itself determined by the human genetic inheritance” (Lerdahl & Jackendoff, 1983a, p. 281).

---

33 See Lerdahl & Jackendoff, 1983a, p. 280.
The hypothesis of innate universals in music perception has been discussed by different authors and in various fields of research.\textsuperscript{34} However, we must be careful not to relate the idea of universals in music perception to any one musical grammar or constrain it only to tonal music, as authors such as Scruton have. At the same time we have to recognize the strong influence of culture and experience in shaping the way listeners understand music. A preferable account concerning musical universals considers the possibility of common cognitive principles (for examples, Gestalt principles) and perceptual structures (temporal structure, intentionality, expectation, learning and memory) which are shared by all people at a very general level. These principles are the basis for every musical experience and for the acquisition of musical knowledge.

**The subjective aspect of music comprehension.**

In discussing features that characterize the basic perceptual understanding of music, it is important to consider the subjective character of the listener’s musical experience. As Clarke observes: “the sounds of the world reach our ears in a very indeterminate raw state, and that individual perceivers then make sense of them according to a whole range of factors: the specific context of the sounds, what the listeners were hearing most recently, differences in how they focus on the sounds, their previous experience or training, and so on – in short a whole variety of “processing” differences largely based on mental representations or memory processes of one sort or another” (2005, p. 11). The man in the...
The man in the metro thus organizes the auditory environment on the basis of his experiences, musical knowledge, attention, concepts and intentions. He has a particular auditory experience which acquires musical meaning because of the way he hears and organizes sounds which have specific acoustic and physical features. This determines his basic perceptual understanding and the meaning of his experience, which might be different from that of the other people in the metro. Someone else, for example, might not recognize the presence of music within the complex and confusing auditory environment of the metro station, or they may have the feeling that those sounds do not represent music because they are unfamiliar or difficult to distinguish and organize.

As McAdams and Bigand observe,

the postulate of a cognitive approach to perception is that the sensory information must be interpreted in order to give rise to a coherent perception. Interpretation is necessary since the information contained in the stimuli that reach the sensory organs is not always sufficient to form a coherent image of the surrounding sound environment. In these cases, the perceptual system must represent and then compare auditory information that is not directly present at the sensory level. [...] When the sensory data that are immediately available are found to be insufficient, the perceptual system analyzes the situation by taking into consideration knowledge that it has acquired of the surrounding sound world (1993, p. 2).

In this way, the man in the metro subjectively organizes and interprets the auditory information he receives from the metro’s acoustic environment.

Some aspects that support the idea of a subjective organization of sounds while listening to music are:

- the influence of learning. This shows how personal experiences determine the way in which a listener organizes the sound environment in musical forms;
- the concepts one has about music. The idea of what a musical genre or form such as sonata form is unconsciously influences the comprehension of a piece of music. Thus, the concepts a listener has in their mind determine their understanding and the way they organize the sound environment;
- the beliefs and intentions one has before listening to a piece of music. The man in the metro, for example, who does not expect music, hears it differently from a person who goes to a concert;
- the cultural and historical context in which a listener lives and, therefore, her point of view in the world shapes her way of listening to music;  
- the way listeners develop their neural and cognitive abilities to process sounds. Because of the plasticity of the neural system and the dynamic character of human cognition, each perceiver develops a subjective way of processing sound information which goes beyond common human structural features such as, for example, temporal structure, specific brain structures and cognitive processes like grouping and spatial localization. The extreme cases of people with brain damage show that the physical structure of the brain influences (or constrains) the way we listen to music (Sacks, 2007).

Concerning the previously mentioned perceiver-environment relationship, the idea that the listener lives in constant interaction with the auditory environment shows that people organize the world subjectively. This does not mean that music – as well as all the other events we can experience in the world – is only a subjective matter. If that were the case, it would be impossible to have any kind of musical communication. As it has been underlined by Clarke and his ecological approach to music, the world and the sound environment have their own structures. People perceptually organize them differently because of the influence of all the factors considered above (culture, context, concepts, beliefs, cognitive and neural structures). What guarantees musical communication among people is the ability to learn certain regularities of a shared environment, as well as the possession of similar perceptual structures and processes that allow individuals to pick up and adapt to new or changing auditory environments. This clearly shows the reciprocity that characterizes the perceiver-environment relationship.

Many cognitive studies on music perception may be criticized because, in the attempt to give a causal account of the way people perceive music, they do not develop the subjective aspect of musical listening. This is firstly due to the fact that a lot of empirical studies focus only on Western tonal music, which presents a model of cause-effect and is a well established musical tradition. If we consider different kinds of music, in particular those that have been developed during the twentieth century and begin to stray from
traditional concepts of tonality (for example, *musique concrete*, free jazz, atonal music and so on), we notice that it becomes difficult to give a causal account of the way sounds are organized and the effects such organization has on the listener’s musical experience.

Subjectivity is important because it allows the constitution of meaningful musical experiences even when listeners experience unfamiliar music. Moreover, it allows music to be open to different interpretations rather than being tied to fixed semantic content. Although people agree that they are listening to music and are therefore able to understand a sequence of sounds as music, they have different experiences of the same musical event, which acquires a specific meaning for each individual.

Intentionality.

Another aspect which is very important for musical comprehension of an event occurring in the auditory environment is the intentional structure of human consciousness. Human consciousness always refers to something, and intentionality is the *directedness* or *aboutness* of the mind. The word ‘intentionality’ is derived from the Latin *intendere*, which means ‘to aim in a particular direction’. Gallagher and Zahavi claim that “intentionality has to do with the directedness or of-ness or aboutness of consciousness, i.e. with the fact that when one perceives or judges or feels or thinks, one’s mental state is about or of something” (2008, p. 109). Thus, intentionality refers to the idea that experience is always directed to some object and that intentional structure characterizes all conscious acts. Intentionality is a ubiquitous character of consciousness and, in the phenomenologists’ opinion, this means that all consciousness (all perceptions, memories, imaginings, judgements, etc.) is *about* or *of something*.\(^{35}\) The intentionality of perception, for example, is described by Gallagher and Zahavi in this way:

---

\(^{35}\) I do not take into account the historical debate about the concept of intentionality. I consider the meaning it acquired in the twentieth’s century phenomenology and interpretation of Husserl’s thought.
When I see a particular object in the street, I see it as my car. Perception is not a simple reception of information; rather, it involves an interpretation, which frequently changes according to context. To see my car as my car already suggests that perception is informed by previous experience, and at least in this sense Locke and the empiricists were correct to suggest that perception is educated by experience. One should think of this as perception enriched by experience and by habitual, as well as customary, ways of experiencing things rather than as a case of perception plus thought. It’s not that I perceive x and then add something quite different and novel, namely the thought that this is my car. The perception is already meaningful (2008, p. 7).

This also indicates that the notion of meaning is central to intentionality and its movement. For example, a mental state that exhibits intentionality can be expressed in the form of a mental representation. In this sense, perception can represent (be about, be directed at) various objects in the physical environment and express how those objects appear in consciousness with a determinate meaning: a particular object in the street, for instance, is represented at the perceptual level as my car.

Intentional states are related to the world in different ways, and an interesting question concerns the way in which intentional states acquire their content. I propose to consider the process of comprehension as having an important role in intentional movement. Its tendency to organize the information a perceiver receives in specific ways is important for the constitution of a particular meaning which is based on her experience. From this perspective, comprehension, as an act of organization, may be considered as a way through which intentionality constitutes a meaningful world. Through comprehension and its movement toward the environment, intentional structure replies to the need of one’s consciousness (assuming that they are always consciousness of something) in order that they are able to grasp the structures of the world and constitute meaningful experiences.

Gallagher and Zahavi start their chapter devoted to intentionality in the following way:

As you read these words you are conscious of something. Maybe you are conscious of the words, or maybe you are mainly conscious of the meaning of the words as your eyes skim over the lines of the text. If you look up you will find yourself in some kind of environment. Perhaps you are sitting with this book at your desk; or maybe you are reading this book as you soak in the tub sipping a glass of champagne (2008, p. 107).
The organizational process and structural composition of human perception is fundamental in order to have this kind of consciousness and immediate awareness of the world. Research in cognitive psychology and neurobiology have contributed to our understanding of how sound organization is realized at the perceptual level and may contribute to the study of the structure of human intentionality.

At the perceptual level, for example, we understand and represent an object as a book or a sequence of three notes as a consonant chord. Consequently, we constitute the meaning of ‘book’ and ‘consonant chord’ in relation to our experience with those objects. This may only occur if we organize the information we receive about the elements that constitute those objects in specific ways during our initial experiences with them. In this way we realize – through the process of comprehension – the intentional movement of our consciousness, which relates the human being to the world and replies to her needs by giving meaning to the environment around her.

Intentional movement is a constitutive part of the perceiver-environment relationship and is at the basis of the immediate awareness of objects as having specific meanings. Gallagher and Zahavi observe that “[o]ne is never conscious of an object simpliciter, one is always conscious of an object in a particular way, be it from a certain perspective, or under a certain conception or particular description” (2008, p. 112). It is in this sense that perception is already meaningful: the subject is conscious of an object as having a specific meaning without involving conscious reflection. The intentional object – which is “not a special kind of object, but rather the answer to the question of what a certain intentional state is about” (2008, p. 114) – is thus constituted in one’s consciousness according to the way its elements have been organized. We can then affirm that the tendency of our consciousness ‘to be about something’ needs a certain level of organization in order to comprehend the environment. Such constitution is a subjective process to the extent that it is influenced by experience. We can explain this by showing how our perception organizes the world around us, and through which mechanisms it realizes meaningful representations and experiences.
There is a wide variability in the way people listen to music and in each individual’s basic musical understanding of sound events occurring in complex auditory environments. Listeners understand music differently and have different comprehension of it, which is the result of their intentions when listening to music. For example, I have an intellectual understanding if I listen to music with the intention of discovering the harmonic relationships that characterize its structure. However, the same piece of music may be understood in different ways because of alternate intentions (which might be determined by the context, recent experiences, knowledge, etc.).

During my research I will focus on the unconscious level of intentionality using music perception as an example. A listener is not conscious of the organizational processes through which she comprehends a sequence of sounds as music. However, in this way she realizes the intentional movement of consciousness and comprehension becomes the fundamental process through which she is conscious of an acoustic experience as having a specific meaning. Without comprehension, a listener cannot be aware of a sequence of three sounds as having the meaning of ‘consonant chord’. In general, without any kind of organization of what she hears, the listener cannot be conscious of music; she cannot perceive her auditory experience of a sequence of sounds as having musical meaning.

Clarke provides an example which illustrates “how the primary function of auditory perception is to discover what sounds are the sound of, and what to do about them: these are the sounds of someone else’s CD” (2005, p. 3). This suggests that when you hear what the sources of particular sounds are, you also acquire better understanding of what those sounds mean. On the basis of these last observations, intentionality and basic perceptual understanding of sounds as music may best be considered by addressing the problem of musical meaning. I will discuss this problem in the next chapter by elaborating on the notion that musical meaning is strictly related to the process of music comprehension and the way listeners subjectively develop such comprehension.
Basic musical understanding.

Comprehension is important in explaining how a listener develops musical experiences at different levels of consciousness. In this chapter, I focused on the main aspects that characterize what I call – following Levinson – ‘basic musical understanding’. Such comprehension is developed at the perceptual level and is fundamental to every kind of musical experience and more complex levels of comprehension. It allows the man in the metro to have meaningful auditory experiences and perceive the acoustic environment not as a confused whole of information, but as an ordered system with a meaning.

Listeners understand the acoustical environment as music at the perceptual level when they organize the sounds they hear according to their experiences, learned knowledge, concepts and beliefs. Two perceptual structures of human consciousness – the temporal and the intentional structure – are fundamental in order to constitute a musical experience, as well as the perceiver-environment relationship that has been underlined both by the ecological approach to music (Clarke, 2005) and the phenomenologist point of view (in particular, Merleau-Ponty). This comprehension is subjective: as we will see, this means that there is no wrong comprehension at the perceptual level, but only different ways of experiencing music. In this sense, the concept of ‘basic musical understanding’ that I propose is different from that proposed by Levinson. Underlining the subjective aspect of the perceptual organization of music, my aim is to argue for the idea that, at this level of comprehension, it is not necessary to grasp the ‘correct’ structure in order to have meaningful auditory experiences. In this way, an unfamiliar sound experience may become musically meaningful. On this basis, ‘basic musical understanding’ can be defined as a perceptual process which organizes the auditory environment and allows for immediate awareness of music. It is determined by learning and intentionality, and it has a strong subjective character. The problem of meaning and what a meaningful musical experience is will be the topic of the next chapter.
THE CONSTITUTION OF SUBJECTIVE MUSICAL EXPERIENCES

Introduction.

Through ‘basic musical understanding’ the man’s auditory experience in the metro acquires a particular meaning. Due to this understanding he is immediately aware of music at the perceptual level and he understands certain acoustic events occurring in the complex auditory environment of the metro station as having musical meaning. His subjective auditory experience acquires a specific meaning in relation to the context and on the basis of the way he perceptually organizes sounds— which is determined by learned knowledge, subjective experiences and the structure of the environment in which he lives.

In this chapter I take into account the concept of ‘constitution’ and relate it to ‘basic musical understanding’ by considering the process through which the man’s experience in the metro acquires subjective meaning. This process allows him to know and become familiar with the world in which he lives and to be immediately aware of his perceptual experiences as having specific meanings.

The problem of musical meaning.

When we take into account the problem of meaning in music, a series of different and important questions arise. Such questions include: what is a musical meaning? To what does it refer? Does music communicate, represent, express or denote particular meanings? How do sounds acquire a musical meaning? These are all well-studied questions which have been the object of various philosophical debates. In addressing the first topic: one type
of proposed musical meaning is the ‘formal meaning’, which has been a prevalent point of view maintained by formalist theorists and supported by a lot of musicologists interested in music analysis. From this perspective, musical listening is a structural type of listening and the meaning of a musical composition is in the music itself. Musical meaning thus corresponds to musical structures and relations such as themes, harmonic progressions and formal prototypes, and from this viewpoint music is considered as a communicative tool because of its syntax (which allows structural meanings to be shared and communicated). Therefore, to reply to the problem of what people hear, formalist theorists such as Kivy (2001; 2002), Hanslick (1986), Lerdahl and Jackendoff (1983a) relate meaning to musical structures, and music comprehension to the understanding of such structures. This presupposes a knowledge of the syntax which is at the basis of a particular musical style, and listeners need a high level of expertise in order to pick up the topical structure of the music they are listening to.

An important representative of the formalist theory is musician and music critic Eduard Hanslick, who claims that musical meaning is inherently musical and any understanding of music’s meaning has to be predicated based on an understanding of its compositional structure. According to Hanslick (1986), the beauty of music itself lies in its structural features, and formal meaning is related to comprehension of the work’s audible musical structure. In Budd’s words: “[t]he core of musical understanding – of hearing music with understanding – is the experience of what I shall call the intramusical meaning of a musical work, that is, the work’s audible musical structure, the musically significant relations (melodic, harmonic, rhythmic, and so on) that obtain amongst the sounds and silences that compose the work” (as cited in Koopman & Davies, 2001, p. 261).

In Meyer’s opinion, meaning is related to the expectations an “Ideal Auditor” creates while listening to music and the emotions they determine. Musical expectations are a function of (and are built on) stylistic norms under which music acquires its meaning. In spite of this and in contrast to the narrow formalist view, music is considered as carrying

---

36 The ‘Ideal Auditor’ is the listener who knows and is very familiar with the style of the musical pieces he hears. From Meyer’s viewpoint, the Ideal Auditor is able to create musical expectations about probable outcomes that are quite likely to be very close to the statistical frequency prediction.
extra-musical meaning which arises because of a reference or analogy with some other events or emotions. This describes music as an intentional object (in the sense that musical events are about something), and maintains that musical meaning is understood as corresponding to the intentional content of a musical event. When music expresses extramusical objects and concepts, we have a ‘designative meaning’, which has often been related to musical emotions. In the paper Meaning and Music, for example, Donald Sherburne defends the view that the meaning of music is a referential meaning that refers to the extra-musical world of emotional states (Sherburne, 1966). According to referentialists such as John Hospers (1946) and Donald Ferguson (1960), musical meaning “lies in the relationship between a musical symbol or sign and the extramusical thing which it designates” (Meyer, 1956, p. 33).37 Such a perspective involves those theories for which music is a representational system.

Music is representational in the sense that it is a symbolic system which carries extramusical content.38 For example, in both Peter Kivy (2001, p. 96) and Susanne Langer’s (1942) view, music embodies emotions as symbols. In Bigand’s opinion music is a form of symbolic expression that actually communicates emotions in a sound form. From these points of view (as for formalists), music is a human construction made up of sounds which are highly organized and embedded in an elaborate system of rules and practise.

Following this perspective, Charles Nussbaum proposes an interesting theory. In studying what musical representations are, he introduces an interpretation of the variety of musical meanings by applying the ‘body-in-the-mind’ theory of metaphorical understanding. With this theory, meaning is generated according to a symbolic system learned through analogical reference to an extra-musical meaning. In the fifth chapter of his

---

37 In the book Emotion and Meaning in Music, Meyer (1956) develops an interesting discussion about the difference between ‘absolutists’ and ‘referentialists’ in relation to the problem of meaning in music. Meyer identifies his own position as ‘formalist–absolute expressionist’: in acknowledging that music can and does express referential (designative) meanings as well as nonreferential ones, he exhibits an eclectic and certainly permissive view. But he has been criticized for failing to make clear the modus operandi of this referential meaning in music.

38 Theorists who follow this perspective investigate the problem of how music may represent anything external to itself in parallel with the rise of semiotics and cognitive studies. These subdisciplines, although entirely distinct from each other, give far greater attention to mental processes than earlier theories which concentrated exclusively on the music itself.

See, for example, Nussbaum (2007); Davies (1983); Langer (1994); Zbikowski (1998).
work \textit{The Musical Representation. Meaning, Ontology, and Emotion}, Nussbaum (2007) addresses the problem of emotions, which are considered extra-musical representations to which music refers and the content of which music expresses. Emotions thus represent a kind of extra-musical semantic content based on a musical syntax. To summarize, Nussbaum relates meaning to a representational rule-governed musical system, in which the meaning of a musical event relies on extramusical content which is based on a specific syntax. Music does not have concepts in itself, although it is conceptually organized.

Another way of discussing the problem of meaning in music is by considering music as socially constructed. In his book \textit{Music, Performance, Meaning}, Nicholas Cook (2007) is concerned with the lack of an adequately theorized conception of how music might support (or not support) the meanings ascribed to it. In his opinion, there is no such thing as ‘purely musical’ meaning. Music supports meanings that are socially and culturally determined. Cultural influence on musical meaning is well expressed by the concept of wisdom presented by Susan McClary in her \textit{Conventional Wisdom. The Content of Musical Form} (2000). According to McClary, conventional structures are ideological formations which are socially, historically and culturally influenced. It is these assumptions which allow cultural activities including music performing, listening and composing to ‘make sense’. They are not fixed and stable, but subject to changes. For example, she considers tonal music, and more specifically the blues genre, as two conventional schemata\footnote{In McClary’s book, conventional schemata express the organizational structures that characterize a musical system and are learned by listeners through exposure. Schema (pl. schemata) is a fundamental concept in music perception to which I referred in the first chapter and I will often return in the paper. A schema, as defined in the cognitive sciences, is a mental structure that represents common qualities of some aspect of the external world. It is the organization of experience in the mind or brain that includes a particular organized way of perceiving cognitively and responding to a complex situation or set of stimuli. In psychology a schema is a pattern imposed on complex reality or experience to assist in explaining it, mediate perception, or guide response. Schemata are perceptual structures which organize our knowledge and assumptions about something and are used for interpreting and processing information. In Kantian epistemology schema is a concept, similar to a universal but limited to phenomenal knowledge, by which an object of knowledge or an idea of pure reason may be apprehended. Schemas influence our attention, as we are more likely to notice things that fit into our schema. If something is contradicting our schema it may be encoded or interpreted as an exception or unique. People use schemata to organize current knowledge and provide a framework for future understanding. As a result of schemata we might act in a way that makes expectations come true. Thus, schemata are an effective tool for understanding the world. Through the use of schemata, most everyday situations do not require effortful processing. In this way, people can quickly organize new} which
reflect rules that are usually acquired by listeners through exposure. In relation to the social aspect of the musical experience, the meaning of particular pieces of music can also be identified with the intentions of the individuals who composed them. In this sense, listeners understand musical meanings when they grasp the composers’ intentions.

Another type of meaning that music can acquire is a literal one. The musical meanings that McClary presents, for example, do not come from the music itself, but are highly related to lyrics. In this way, she follows a general tendency to assimilate meaning through verbal signification.

A further issue in the debate about musical meaning is related to meaning and the communicative aspect of music. If music is indeed communicative, what precisely does it communicate? If we agree in saying that music has meaning, it may also be said that music is a form of communication. In this way we allow that through music, we can communicate something meaningful as it happens just as we may with language. If we disagree with a representational theory based on a syntactical system, the problem of what music communicates becomes more complex because unlike language, its semantic content is not fixed.

As it emerges from this short overview about some of the existing theories concerning the problem of meaning in music, we note that music can acquire very different meanings in relation to the context under which it is experienced. My aim is not to discuss what a musical meaning is or what kind of meanings music expresses. In order to understand how the man in the metro’s auditory experience acquires a specific musical meaning (for example, ‘there is someone who is playing in the metro’, ‘this is a famous Christmas song’, ‘there is someone who is playing the violin’), I focus on the processes through which sounds that are perceived in a complex auditory environment become the expression of a musical meaning for the subject.

In relation to the problem of music comprehension, I try to highlight the main aspects of the process of constitution of meaning. As noted, comprehension is the condition under which the listener may be aware of her perceptual experiences as having meaning. In

perceptions into schemata and act effectively without effort.
the next chapter, I will present some cognitive studies that provide evidence for the way such a process is developed at the perceptual level, in order to support its relevance for the immediate understanding of the acoustic environment as musically meaningful. In this way, I will assert that, even if the most recurrent structural features of a musical style are unconsciously learned by listeners and they influence to some extent the meaning their musical experience acquires, such a meaning is not in the music itself, nor it is constrained by syntactical features in a well defined semantic field. There is an ineliminable subjective aspect in the way sounds are organized and a musical meaning constituted.

Music perception is a dynamic and intentional process that allows listeners to constitute musical meanings even in relation to unfamiliar musical events for which they do not know the musical syntax (if there is any). They organize sounds and constitute meanings based on past experiences, learned schemata and particular cognitive processes such as expectation. This is fundamental to the comprehension of the auditory environment, to the listener's familiarity with it and to the fact that the listener is always consciousness of something. I will further develop this idea through the arguments I will present in this and the coming chapters.

As we have seen, the listeners' experience of 'hearing sounds as' music indicates the intentional movement of human consciousness. The content and the meaning a listener acquires during a sound experience is partly dependent on this intentional movement, which characterizes 'basic musical understanding' and the way it is realized. The organization of sounds into structural forms reflects the idea that consciousness must go beyond random acoustic events; one must always be consciousness of something. In this way the listener meaningfully understands the auditory environment.

The meaning of music is not in the music itself, but is constructed by the listeners through their relationship with the environment. This underlines the experiential, subjective aspect of musical listening. The role of the perceiver in the determination of such a meaning is therefore very important, and this role highlights the subjective aspect of the musical experience. Unlike language, music does not have a fixed semantic content.
A piece of music should be conceived as an indefinitely extended series of auditory traces from which meaning emerges through an act of organization and with reference to extra-musical content. There is a complex articulation between music and its meanings which is influenced by culture and subjective experiences. As Cook observes, music “has the potential for specific meanings to emerge under specific circumstances” (Cook, 2007, p. 224-225). In the light of this, the same piece of music may support a plurality of meanings, which are highly dependent on culture and subjectivity. Because of this plurality of meanings, Cook defines a difference between potential meaning and actualized meaning. Based on an individual’s interpretation, indefinite potential meanings are transformed into actualized meanings, which are constrained by cultural influence and dependent on subjective experiences. Cook writes: “as I have suggested, musical meanings are actualized through processes of critical interpretation that are culturally and historically contingent; in this sense meaning is indeed a cultural construction” (2007, p. 230-231).

On the basis the theories analysed thus far and in the framework of the present work, I argue that music does not have meaning in itself, but rather acquires meaning and becomes expressive of meaning because of a consciousness that allows one to perceive and understand sound events occurring in the environment in a particular way. Emotions, for example, may be considered as a type of meaning that the musical experience acquires. That is, emotions are not in the music, but arise from listening to it because of the way sounds are processed by the neural and cognitive system of the listener. In such a process, expectation plays a very important and determinant role.

A sequence of sounds thus acquires meaning because of the activity of the human mind that relates sound structures to particular meaningful experiences. The organization of sound into structural forms is fundamental to the musical experience and its meaning, but the experience is not reducible to structural organization alone. Making reference to some phenomenological concepts, I focus on the way (not in the why) in which a perceiver constitutes meanings during musical listening. Moreover, I consider what happens at the
cognitive level during the perception of music. Meaning always depends upon the way the perceiver listens to music, and how she processes the sounds she hears.

**The constitution of a meaningful musical experience.**

**Introduction**

In the preceding pages I have briefly discussed the ways in which musical meaning is understood by some theorists. For the purposes of the present work, however, I do not focus on the content or the object of musical meaning, but rather on the process through which a musical meaning is constituted in consciousness. With reference to the example I introduced at the beginning, I try to understand the way the man in the metro becomes aware of his auditory experience as having a musical meaning. This might be helpful in order to better understand the concept of ‘basic musical understanding’ which I discussed in the previous chapter. In particular, through the concept of ‘constitution of meaning’ I try to understand what it means to have perceptual awareness of an auditory experience as musically meaningful. From a phenomenological perspective, the human perceptual experience – in its encounter with and action in the world – is already meaningful in the sense that the perceiver does not need conscious reflection in order to understand the environment as having specific meanings. Analyzing its features, Gallagher and Zahavi describe perception not as a simple reception of information, but as a process which involves interpretations that may change according to the context and may be enriched by experience. This idea underlines the experiential aspect of musical listening, the immediate relationship a subject has with the sound environment and the way she understands it. In other words, this describes a comprehension which the listener has at the perceptual level without the mediation of conscious thought. Interesting in this respect is the idea of

---

40 In relation to the problems of how music incites the emotions, Juslin and Västjäkl (2008) argue that to understand the musical experience we have to take into account the underlying mechanisms which allow listeners to hear music.
‘experiential formal meaning’, which is well expressed by Davies and Koopman in the following quotation:

the fact that we can explain music in terms of reasons provides a sufficient justification for speaking of musical meaning. However, there is a more fundamental sense in which music can be said to have meaning. To understand music as meaningful, it is not necessary that we can explain the progression of the music. Meaning can be understood immediately in the musical encounter, without reasoning at all. Music appears to have an experiential, rather than a discursive, logic. We do not merely perceive a succession of patterns in music. Instead, we experience the musical parts as connected into a dynamic whole. There is sense to the way music progresses. Music presents itself as a continuous process in which, at every moment, what we hear follows in a compelling way from what came before; that is, music proceeds not as the temporal succession of otherwise unconnected elements but as the unfolding of an integrated whole (2001, p. 264).

An important aspect of such ‘experiential formal meaning’ is that it is the result of a dynamic process developing in time. I will consider this in the sixth chapter through analysis of the temporal structure of human consciousness. Moreover, the succession of temporal moments into an integrated whole introduces the discussion about expectation and gives a conceptual framework within which I may discuss its role in the constitution of meaning.

The concept of ‘constitution of meaning’ and the immediate awareness of music that it allows at the perceptual level is strictly related to the idea of ‘basic musical understanding’. Listeners are aware of music at the perceptual level when they perceive sounds as organized into structures which are recognized as musically meaningful based on experience, culture and expectation. In order to understand this point, I try to analyse the features that characterize the process of constitution and its development in the case of music perception. I begin by referencing aspects of the process of comprehension which I have previously outlined, as these aspects are fundamental to an understanding of how a perceiver organizes their auditory environment into meaningful forms. They reflect a natural process of relating to the environment and living in the world. ‘Natural’ here means that listeners have a spontaneous tendency to develop specific ways of organizing the auditory environment which is not culturally determined. For example, learning the regularities of the environment and organizing sounds based on these regularities reflects a

41 ‘Natural’ in this case means not arbitrarily determined by cultural constraints.
natural tendency and a spontaneous, immediate way to understand the world. In this way, listeners adapt their cognitive structures to the features of the environment. At the same time, they shape their way of hearing music, which becomes subjective since it is based on the experience and knowledge of the listener.

The specific perceptual organization of sounds into meaningful forms which a subject develops is thus shaped by learning, cultural knowledge, concepts and beliefs, and, in general, determined by the subject’s point of view in the world. At the end of the chapter I will introduce the mental process of expectation as having a fundamental role in shaping the subjective musical experience and constituting perceptual musical understanding of the auditory environment. Prior to discussing the process of expectation, however, I will present a phenomenological account of the concept of ‘constitution’ of meaning and explain why I use this concept instead of that of ‘attribution’ of meaning.

The concept of ‘constitution’.

The man’s perceptual awareness of music in the metro station may be described as a process of constitution of meaning reflecting his subjective comprehension of the auditory environment. The work The Formation of Husserl’s Concept of Constitution by Robert Sokolowski (1964a) well explains the concept of ‘constitution’.42 He takes into account the formation of the concept itself within the whole of 19th Century philosopher Edmund Husserl’s work and in relation to individual subjectivity.

Sokolowski also underlines the way in which the concept of constitution relates to the concept of intentionality43 and the problem of inner temporality – two concepts which, as we have seen, are fundamental to both musical understanding and meaning. One of the

42 Husserl’s choice of the term ‘constitution’ comes from the influence that the neo-Kantian Paul Gerhard Natorp had on his philosophy: “[t]he general sense of the term in Natorp is the constitution of objects of encounter by the application of certain subjective categories or apriori laws to immediately given sense data. Constitution of objects of encounter takes place in subjectivity. […] Constitution is thus the process by which subjectivity forms objectivity by virtue of its own activity; Natorp often speaks of the objectivating function of consciousness” (Sokolowski, 1964a, pp. 214-215).

43 Sokolowski notes that most of Husserl’s treatment of ‘constitution’ in Investigations consists of structural analyses of intentionality.
main questions at the basis of Husserl’s work is: where does meaning come from? He explained the origin of meanings and objects through the concept of constitution by arguing that it represents an articulated process of consciousness that makes things present to us. Constitution governs the way meanings come to be and is related to studies of the sources of primitive concepts and forms. Within this framework particular relevance is given to sensory perception, which has the function of constituting the world in the mode of presence. Moreover, the explanation of the constitution of immanent objects and meanings is an explanation of how objects of the world are experienced in consciousness. In order to develop his general theory about the constitution of intentionality in connection with the constitution of meaning, Husserl (in his *Logical Investigation*, 1900; 1901) addresses a problem concerning the relationship between the subjectivity of knowing and the objectivity of the content of knowledge, and the way objectivity arises from subjectivity. For example, through the constitution of musical meaning, the listener tries to objectify a subjective sound experience. The listener experiences an acoustic event, which she subjectively interprets and perceives as a musical event. An experienced event is presented to the consciousness as an object having a precise meaning: that of being ‘music’. The event or object of the world is subjectively known and experienced, but it is considered an objective content because it transcends immanent consciousness.

Husserl also makes a distinction between meaning and the act of meaning. Such a distinction is based on the fact that one identical meaning can be repeated in several numerically distinct acts of meanings. As Sokolowski observes, “the objectivity of meanings is ideal, not real. Meanings have their existence only within human acts, but one and the same meaning can be realized in a multitude of numerically distinct acts” (1964a, p. 44). This underlines the idea that constitution of meaning is a process through which human beings perceive music in a subjective way.

Another problem to be considered on the topic of musical meaning concerns the relationship between meaning and the object of reference. Sokolowski writes: “[a]lthough

\[See McKenna (1989, p. 188).

\[In the attempt to describe how subjectivity ‘constitutes’ objectivity, the basic idea is that objectivity comes into one’s consciousness when sensory data are ‘apprehended’ by intentionality.\]
meaning is distinct from the object meant, there is, claims Husserl, a close liaison between the two. […] The structure of an act of meaning is such that it involves both a meaning and a reference to an object” (1964a, p. 45). Moreover, “[w]hen a meaning is constituted, reference to an object is also necessarily established. Husserl uses the term “constitution” also to name the way in which an object is so established, in which it becomes the object of reference for a meaningful intentional act” (1964a, p. 54). In the case of music, the problem concerns what kind of object music is or refers to.

Based on the previously discussed notion of intentional movement, musical meaning is related to an object and indicates how meaning is acquired by the object, thus specifying the point of view from which a subject refers to the object. However, the object never coincides with the meaning. This opens the door to a plurality of interpretations of the same object and, in the case of music, allows the listener to have subjective meaningful experiences of both familiar and unfamiliar musical objects.

Intentionality and the perceiver-environment relationship are fundamental to a discussion concerning the interaction between musical meaning and object of reference. The problem of objects of reference indeed leads to the concept of intentionality. Every conscious act, according to Husserl, refers to an object; that is, every act is intentional by nature. Consciousness is not a self-enclosed island, but essentially involves reference to an object, as it is always consciousness of something. In this way, Husserl excludes the concept of a consciousness that is locked within itself and knows only its own immanent elements or states. In this sense, intentionality relates consciousness to objects through real contact with them, thus allowing interaction between the subject and the world. With reference to the musical experience, this relates consciousness to the physical and acoustic aspects of musical sounds.

The relevance given to the acts of consciousness does not imply that the subject is an entity separated from the world. Contrarily, it is emphasised that sensory data plays a role in the subject’s constitution of meaning:46 every intentional act is performed in

---

46 The distinction in Husserl’s phenomenology is that sensory data is experienced in consciousness, whereas objects are perceived or appear to consciousness through perception. I will briefly consider the distinction between ‘experiencing’ and ‘perceiving’ in the sixth chapter.
association with sensations and is directed towards its object through such sensory data. The objectivating act that allows the constitution of meaning establishes an intentional relationship between consciousness and an object of the world. The object then acquires a meaning and becomes an object of consciousness. This is the basis for every intentional act (perceiving, remembering, evaluating, imagining).

With reference to the musical experience, through the process of constitution, listeners understand sequences of sounds as music and they organize them into musical forms. The sounds that the man perceives in the metro – which are characterized by specific temporal and acoustic features – become an object of consciousness having a particular structure and carrying a musical meaning.

Husserl makes an important distinction between ‘acts of meaning’ and ‘acts of perceiving’, although the former acts are oriented toward perceiving and receive their fulfilment through perception. I argue that meaning is the result of an intentional act (that is, being conscious of) which is developed during the perceptual experience. As a kind of intentional act, perception realizes the constitution of meaning as an experiential and objectivating act and guarantees a relationship with the world. Following this idea, I agree with Sokolowski’s critique of Husserl when he writes: “He [Husserl] declares that meanings arise when sensations are objectivated by an intention; this serves to locate meanings in the structure of intentionality, but it does nothing to explain how they arise in that structure. Meanings are simply there, but their origins are not explained” (1964a, p. 59). A solution to this problem could be found through the involvement of perception. Solokowski proposes that perception is “the logical place to look for a further explanation of the origins of meanings, but Husserl does not investigate it with this problem in mind. When he treats perception, it is only to show how meanings (which are already constituted in an intention) can be recognized or filled in a real object. He does not raise the question of how they can be discovered in perception” (1964a, p. 59). Husserl locates meanings within intentional acts, but this does not show how such meanings arise and how they are constituted within intentionality. Solokowski suggests that a remedy for these difficulties might be found in showing how meanings arise from perception. He observes that
Husserl’s entire theory of perception is elaborated to show how meanings find their fulfilment in acts of perceiving. He always supposes that meanings are already constituted and shows how perception can, in various degrees of completeness, “fill” the meanings that are otherwise only intended. The argument always goes from meanings to their fulfilment in perception; a study in the opposite direction, from perception to the formation of meanings, would correct this weakness in the explanation of intentional constitution (1964a, p. 73).

In his *Investigation* (1900; 1901), Husserl also shows the influence of sensory data on the constitution of meaning, admitting that intentionality is not completely free in forming meanings during acts of perception. This shows the relevance of perception not only for the fulfilment but also for the constitution of meaning, and the fundamental link with the sensory world.

On the level of sensory experience, perception of reality is directed to objects outside of the stream of consciousness. I consider perception as the necessary link to the world: it is responsible for human experience in the world and allows the constitution of meaning. It is through this constitution that sensory data becomes objects of inner consciousness and the human being knows the world. By analyzing studies of the cognitive processes involved in the perceptual organization of sounds, we may better understand some of the possible descriptions for how meanings are constituted during the perceptual act. For this reason, I take into account the underlying mechanisms which allow the experience of a sequence of sounds to acquire musical meaning at the level of perceptual awareness. In particular, I focus on expectation as a determinant process for the constitution of musical meanings.

An important aspect of the process of constitution – understood as an intentional process – is its temporal unfolding. As we will see in detail in the sixth chapter, simple

---

47 Husserl’s studies of reason and reality in *Ideas* (1913) will partially satisfy this need.
48 As Sokolowski observes, “even in the *Investigations*, some elements towards a study of this sort are provided in what Husserl says about the function of sensory contents in the determination of meaning. Intentionality is not completely free in forming meanings in acts of perception. The sensory data it animates impose certain limits on which meanings can be constituted, and it seems that if we are to look for a more fundamental explanation of intentional constitution, it would be found in the sphere of sensibility. Sensory contents could be to simple constitution what simple intentionality is to categorical constitution: the more rudimentary level of experience where an explanation of constitution is to be found. Husserl’s later works do investigate sensations with this in mind” (1964a, p. 73)
sensations that are at the basis of the constitution of meaning are related (synthesized) into a whole through temporal flow. The perceptual apprehension itself is constituted by immanent temporality.\textsuperscript{49} From this perspective, meanings are constituted and experienced through a \textit{manifold} of temporal phases. The temporal extension of musical objects is important for individuality and unity, which they must have in order to be experienced. The flowing presence of time is an essential part of consciousness and the temporal movement is fundamental for the constitution of meaning.\textsuperscript{50} In particular, through constitution and the mechanisms of retention, primal impression and protention, musical objects are brought into existence.

Another important aspect for both the determination of meaning and the understanding of sounds as music is subjectivity. Sokolowski writes: “[s]ubjectivity explains reality; this means that reality does not have an ultimate explanation in itself, but that it must somehow depend, phenomenologically, on consciousness” (1964a, p. 108).\textsuperscript{51} This supports the subjective aspect of musical experience and allows for different interpretations, perceptual experiences, comprehension and meanings of a sound event. Reality (of which the musical experience is part) depends on subjectivity for its sense, and subjectivity is the basis for sense or meaning. From this perspective, the meaning of reality is the result of the subjective process of constitution: “[t]he means by which reality acquires its relativity, the process by which it acquires the sense that it has, is called constitution” (Sokolowski, 1964a, p. 130).

Even if we accept the assumption that reality receives its sense from subjectivity, it does not mean that the subject is separated from reality or that the real world does not have its own existence. In Husserl’s opinion sensations or appearances are not signs or symbols of real things which are not accessible in themselves. The subject does not see appearances

\textsuperscript{49} Immanent temporality is the primal constitutive source of human experience. For an introduction to the concept and its relation to constitution see, for example, Larrabee (1989); Sokolowski (1964b).

\textsuperscript{50} A problem of Husserl’s explanation of temporal constitution and the way sensations are experienced is that it does not account for their qualitative elements, nor does it show why or how one sensation is ‘green’ while another is ‘red’, and so on. As noted by Sokolowski, “we have stressed that Husserl’s temporal analyses are formal. They explain the temporality of subjectivity and the time in which all transcendent objects are located, but they do not explain the material content of objects in time” (1964a, p. 114).

\textsuperscript{51} A relevant role is given to subjectivity by Husserl in \textit{Ideas} (1913), where subjectivity is considered as a region of adequate experience capable of supporting a rigorous science.
and then infer reality from them, but she perceives reality directly through appearances and sensations. The world she perceives and constitutes is the real world; it is reality itself. Under this perspective, Husserl does not identify reality and consciousness, even though they are strictly related: meanings indeed arise from the relationship between world and subject. This is an important point to note in order that we may better understand the relevance of the perceiver-environment relationship in determining the constitution of meaning and basic musical understanding of acoustic events occurring in the auditory environment.

As Sokolowski observes, “[w]hat reality is in itself can be reached by consciousness even though it must remain, in principle, radically distinct from and transcendent to consciousness. This is the mystery of intentionality, the mystery of consciousness” (1964a, p. 135). In this way, he tries to highlight the essential role of this relationship for the constitution of meaning: “real things, since they are units of sense, presuppose subjectivity as their necessary correlate, because sense can only arise in connection with intentional subjectivity” (1964a, p. 137). Consciousness is a necessary condition for the actualization of reality as possessing sense, but it is not the cause of reality. More precisely, Sokolowski adds that “subjectivity does not cause or create senses and objects. It merely allows them to come about. It is their condition, and not their cause” (1964a, p. 159) and “sensations are the raw material which serves as the basis for the constitution of objects and meanings” (1964a, p. 210). For the constitution of meaning both subjectivity and reality are therefore fundamental.52

According to Sokolowski, “[t]he real cannot be conceived apart from its manifestation to subjectivity, but consciousness cannot be conceived without its tendency to the real, and both dimensions must be retained in giving the philosophical basis for constitution” (1964a, p. 219). In light of this, constitution is conceived as the product of the

52 Consciousness is not a sufficient condition for the constitution of meanings. There is a facticity and givenness in the actual content of meaning which is found in objects when they come to be. As Merleau-Ponty also claims in the explanation of human perception, the world is the background of our knowledge. “The existence of the world as a transcendent horizon for our experience and encounter is apodictic; it is the background against which all individual instances of knowing take place, and although we may be misled in certain cases of knowledge, we cannot be misled about the background of knowledge itself” (Sokolowski, 1964a, p. 219).
dialectic between subjectivity and the real. In this dialogical relation, intentional movement plays a determinant role. Indeed, intentionality is essentially orientated toward discovering the real world, which becomes the necessary correlate to consciousness. In this way, reality gives consciousness its sense and consciousness cannot be understood apart from its orientation to the world.

Here we find the enigma of constitution: intentionality constitutes an object that transcends consciousness, which acquires an existence and a sense that becomes independent of subjectivity. At the same time, subjectivity is the necessary condition for the emergence of its transcendent sense. However, in this way it is clear that Husserl’s concept of constitution supposes two elements. The first such element is that subjectivity is a condition under which constitution becomes possible, and the second assumes a certain facticity in what is actually constituted. Sokolowski critiques the fact that Husserl elaborates only on subjectivity. The conditions under which reality emerges in consciousness through the process of constitution must be investigated.

In the case of music perception, meaning depends on the way human beings hear and organize sounds. From a phenomenological point of view, ‘music’ is a meaning constituted by human beings. Music exists in the world in the form of acoustic vibrations, but it acquires the meaning of ‘music’ the moment a subject experiences and perceptually organizes such physical events. Thus, human consciousness is the condition for the emergence of an event in the world as having a specific meaning – that of being ‘music’. According to Sokolowski, meaning is not given by the sensations that are experienced by a subject in her relationship with reality, but by the interpretative act of grasping them. The ultimate grounds for musical meaning is the subject. As Steve Larson observes, “listening to music is a creative process in which we shape the sounds we hear into expressive meanings tempered by our nature and experience” (1997, p. 321).

53 The dialogical relationship between subjectivity and reality gives rise to what has been called the ‘enigma of transcendence’: “Husserl maintains the transcendence of the world in relation to subjectivity. Reality is not reduced to consciousness; it remains transcendent and separate from the mind, and yet, in this very transcendence, it is accessible to consciousness” (Sokolowski, 1964a, p. 165).
The constitution of subjective musical experiences.

In the previous paragraph I presented Husserl’s concept of constitution in order to better understand how an acoustic event is understood as music and how musical meaning arises. I chose the concept of constitution rather than the idea of giving a musical meaning to a sound event, which presupposes that a subject imposes her perceptual (and mental) structures to the environment without any direct relation with the living world.

I turn now to the study of constitution with reference to the specific case of music using the man in the metro as an example. I will consider the processes through which he becomes aware of a musical event occurring in the complex auditory environment of the metro, and how this awareness is constituted as a subjective musical experience.

If ‘music’ is the meaning which the man in the metro’s auditory experience acquires, I consider constitution to be an intentional act of objectivation which is realized during his encounter with the acoustic event thanks to the perceptual structures and cognitive principles involved. During the perceptual experience of sounds coming from the acoustical environment, an individual encounters the world in which she lives. The objects of reference in the world (in the case of music, acoustic events) come to be organized by the subject’s inner consciousness into musical forms which transcend the subjectivity itself. A musical meaning is thus constituted. An acoustic event occurring in the environment comes to be conceived as music; as having the meaning of music. This meaning reflects the intentional movement of the subject and her comprehension of the world. It is transcendent because of an act of objectivation that depends on the subject and is realized because of its relationship with the environment and the way she organizes it.

I will use the concept of Encounter (Erfahrung, used by Husserl to name the consciousness of what is transcendent to subjectivity) as part of the Experience of a subject in the world (Erlebnis, which expresses the consciousness of what is immanent to subjectivity). An experience involves both immanent subjectivity (which is the basis for the constitution of musical objects) and the encounter with the world. It allows the
consciousness of transcendent objects, which represent the meaning that the perceptual experience acquires after an intentional objectivating act. When I use the expression ‘meaningful musical experience’ or, more simply, ‘musical experience’ I am referring to the subjective perceptual act of constitution involving both the condition of subjectivity and the facticity in what is being constituted.

In the case of music, the listener’s perceptual experience of a sequence of sounds occurring in the acoustical environment acquires the meaning of ‘music’ due to an intentional objectivating act that is realized by her subjective consciousness. The expression ‘constitution of a musical experience’ therefore indicates, at its basic level, the comprehension of a sequence of sounds as music. It describes the way in which the listener develops her intentional movement toward the sound environment and constitutes musical meanings during the act of perceiving sounds. In this way, she becomes aware of her auditory experience as having a specific meaning.

As we have previously seen, because of the intentionality of human consciousness, the human experience generally needs to be meaningful. In listening to music, listeners must be conscious of something. In the same way, the man in the metro needs to be conscious of what is happening in the auditory environment in order to act and move adequately. If he is not conscious of someone on the speaker announcing the next station he may not get off at the correct station. If he is not able to recognize that a woman is screaming on the top of the stairs, he will not be aware of possible danger.

An auditory event occurring in the environment is present to the human consciousness, but it becomes a musical event because of the process of constitution. In this way the man in the metro develops a musical understanding through representation of particular auditory events as musical forms. If we consider the constitution of musical meaning to be a kind of comprehension of an acoustic event, then some of its relevant aspects are: the role of subjectivity, temporal development, dynamic intentional movement, the perceiver-environment relationship and the process of learning.

In studying the perceptual experience of music, a lot of music theorists such as Meyer, Lerdahl and Jackendoff have focused only on tonality. This presents a problem in
that such theories follow a model of cause-and-effect that overlooks the extent to which
music is heard subjectively. As Thomas Nagel (1974), who addressed the question ‘what is
it like to be a bat?’, Clarke asks “what is it like to listen to the first movement of
Beethoven’s String Quartet in A minor, Op. 132?” (2005, p. 156). Both highlight the
impossibility of ever knowing what the subjective experience of another organism might be
like. This means not only that there is an irreducible subjective aspect to the musical
experience which we cannot explain nor objectify, but also that a piece of music or a
sequence of sounds can be experienced in different ways and have different meanings in
relation to the way it is experienced. The man in the metro, for example, perceives sounds
differently from the other people around him. Someone else may not even recognize the
presence of music, or will perceive it differently from the man of our example. As Lerdahl
and Jackendoff note, it is rare that two people hear a piece of music in precisely the same
way or with the same degree of richness.

To underline the subjective aspect of the musical experience, we can use the concept
of ‘meaning-for-the-subject’ as presented by Koopman and Davies: “it has to do with the
place something takes in the individual’s life or consciousness, with the specific way she or
he experiences it, and with how this relates to her or his perceptions, feelings, thoughts, and
desires” (2001, p. 268). It is because of the presence of the subject that a musical meaning
arises. Otherwise, a number of physical vibrations would be only a random sequence of
acoustic events. Indeed, in the moment they are perceived by the listener, composed by the
composer or performed by the musician, they become organized sounds. That is, auditory
events become music because there is a subject who organizes and constitutes them as
music. In other words, music in itself does not have any meaning: it is only a series of
acoustic events. Such events become music only when they are organized by a subject in
one way or another. This does not mean, however, that the musical experience is
completely dependant on the listener. The relationship with the environment is
fundamental, as are the music’s acoustic and structural features.

Considering the role of subjectivity in the constitution of meaning, it is evident that
intentions, concepts and beliefs also have an important role. For example, during Christmas
time, the man in the metro expects that in all public places there will be objects that refer to Christmas. As soon as he identifies music, he expects Christmas songs and his musical experience is highly influenced by such expectation.

With reference to the ecological approach proposed by Clarke and his idea of *resonance* with the environment, the relevance of the relationship between perceiver and environment for the constitution of meaning is clear. “Perception is the awareness of, and continuous adaptation to, the environment, and, on the basis of that general definition, the perception of musical meaning is therefore the awareness of meaning in music while listening to it” (Clarke 2005, p. 4). This also underlines the dynamic aspect of perception; its ability to adapt to the changes of the environment while continuing to constitute meaningful experiences. The idea of resonance pertains to a musical understanding for which meaning does not entirely depend upon a subjective interpretation of the environment. In this way, sounds occurring in the acoustic environment are linked to objects of the world and the listener is strictly related to the reality in which she lives.

The perceiver-environment relationship, which is realized through the human body,\textsuperscript{54} is therefore fundamental to one’s representation of the world. As Nussbaum observes, “for the cognitivist, the human mind-brain is an organically embodied representational system, a system that enters into states that are systematically interrelated and that stand in intentional relations to the environment in which the human organism is embedded. […] The mind-brain also represents its environment perceptually by means of the body” (2007, p. xi). From this perspective, the musical meaning that a sequence of sounds acquires during the perceptual experience is not in the sounds themselves nor is it in the subject, but rather emerges from their interaction. This is possible because of the dialogical relationship between perceiver and environment, thanks to which a sequence of sounds comes to be music, to have the meaning of music.

The temporal structure of human consciousness underlines the dynamic aspect of the constitution of meaning. In particular, the ‘experiential formal meaning’ reflects an experience that unfolds in real time. It involves the perception of an ongoing musical

\textsuperscript{54} More specifically, the biophysiological mechanisms involved in sensory perception of the world.
process, not the perception of a sequence of isolated events. As we will see in the sixth chapter, the study of the temporal development of musical experience is helpful in order to better understand how music is constituted in consciousness during the act of perception.

As previously alluded, the experience of particular acoustic events becomes musically meaningful largely due to the influence of learned knowledge which is culturally and socially determined. Listeners learn to perceive sounds as musically meaningful and having specific meanings because they unconsciously internalize the recurrent structures of the music to which they are exposed. In this way, learning may be seen as an appropriate biological strategy which focuses particular attention and sensibility on the reoccurring environmental events.

Musical understanding is often related to comprehension of the meaning a musical event has within a specific culture. Here I do not focus on individual culture-specific meanings, but nevertheless wish to underline the fact that culture influences the way listeners constitute musical experiences at the perceptual level. We have seen in other works (like those of Lerdahl, Jackendoff, Bigand) that simple exposure to the music of a particular culture influences the way music is organized and understood. The man in the metro constitutes his meaningful experience on the basis of his cultural influences. This is the reason why, for example, he may not be able to immediately identify and recognize the presence of music belonging to a completely unfamiliar culture.

Listeners therefore represent the environment in musical forms on the basis of learned knowledge and experiences. The cultural influence also indicates that the musical experience is not static, but refers to social, historically contingent contents. In this sense, the musical experience represents a dynamic process which is open to the environment.

To better understand how musical meanings are constituted in consciousness during the perceptual experience of the man in the metro we should take into account the cognitive and neural processing of sound information that is involved. Such processing is fundamental to the comprehension of sounds as music and to the constitution of a perceptual awareness of music within a complex acoustic environment.

For an example see Clarke (2005, chap. 2), where he proposes an analyses of Jimi Hendrix’s performance of ‘The Star Spangled Banner’ at the Woodstock festival in 1969.
aforementioned *Gestalt* psychology, for example, is clear in showing that understanding the sound environment as music is a matter of grouping stimuli into patterns and relating these patterns to one another. Some relevant cognitive principles involved in music perception are those of similarity, grouping and sound localization. For our purposes, in addition to the principles proposed by the Gestalt psychologists, I will focus on the role of another cognitive process: expectation.

After analyzing characteristics of the constitution of meaning, I propose to relate constitution to the idea of ‘basic perceptual understanding’ which I discussed in the previous chapter. Through this process, the perceiver’s consciousness represents the world in a meaningful way: the listener perceptually understands the acoustical environment as having a musical meaning.

Comprehension is the process through which a sequence of sounds is organized in order for the listener’s perceptual experience to acquire a musical meaning. Could the explanation of the process of constitution of musical experiences (i.e., experiences having the meaning of music) give a new path from which to discuss the problem of meaning with respect to unfamiliar music? From a formalist point of view, music that is completely unfamiliar to listeners is meaningless. From the experiential point of view, every perceptual experience is already meaningful. The moment I am aware of listening to music, I can say that I am having a meaningful auditory experience, even if I am not familiar with the music being heard.

The meaning of an acoustic event and its perceptual comprehension is usually dependent on the cultural and social environment in which it is experienced. As Clarke observes:

> ...in ecological theory, perception and meaning are closely related. When people perceive what is happening around them, they are trying to understand, and adapt to, what is going on. In this sense

---

56 In this sense, we do not have to make the error of the atomists, who attempted to explain and understand music as a succession of separable, discrete sounds and sound complexes. As we will see in detail in the sixth chapter, the temporal structure of human consciousness is fundamental in order to create relationships between successive sound events.

57 As we will see in detail, sound localization is a listener's ability to identify the location or origin of a detected sound in distance and direction.
they are engaged with the meanings of the events in their environment. As I explore further in the next chapter, to hear a sound and recognize what it is (for example the sound of the mail being delivered through the letterbox) is to understand its perceptual meaning, which will result in corresponding actions. By contrast, to hear a sound and not recognize what it is, is to fail to understand its meanings and thus to act appropriately (2005, p. 6-7).\textsuperscript{58}

What then happens in the case of unfamiliar sounds? Do we not have comprehension at all? I propose that in such a case we are able to have meaningful experiences and that the process of expectation has a great influence on the constitution of such meanings.

\textbf{Introduction to expectation.}

I have previously introduced the idea that the way in which an auditory experience is constituted as musically meaningful could be investigated by considering cognitive studies which focus their attention on music perception. The cognitive principles that are unconsciously used by a perceiver to organize sounds into musical forms help to explain how the man in the metro constitutes his experience and is immediately aware of music in the complex auditory environment of the metro station. As mentioned, I will focus my attention on one of these principles: the process of expectation. In particular, I will discuss its role and influence on the constitution of subjective musical experiences.

The basic idea concerning musical expectation is that through temporal movement, which is developed during one’s own experience and through the intentional act that guides such development, a perceiver builds a complex system of expectations that influence the meaning of her experience. As Titchener and Broyles observe, “if information, expectation, and feedback are necessary at the first stages, they are also necessary for determinate meaning, which obviously presupposes the earlier stages in the organically developing understanding of the piece. This development is obviously a temporal one” (1973, p. 25). Focusing on the ‘process’ of constitution instead of the meaning itself, expectation process plays a relevant role in the constitution of meaning. In Vuust and Frith’s opinion, “musical

\textsuperscript{58} Clarke is interested in the perceptual meaning of sounds and underlines the relationship between this perceptual meaning and musical understanding.
expectation is a good candidate for the fundamental mechanism guiding the experience of musical meaning as well as emotion” (2008, p. 600). In particular,

…most music theoreticians consider musical anticipation as one of the principal means by which music conveys meaning and emotion. According to this point of view, understanding music is related to the anticipatory interplay between local auditory events and a deeper structural layer partly inherent in the music itself, and partly provided by mental structures in the listeners that is induced by music. In short, the musical experience is dependent on the structures of the actual music, as well as on the expectations of the interpreting brain. These expectations are dependent on long-term learning of musical structures (culture-dependent statistical learning), familiarity with a particular piece of music, and short-term memory for the immediate musical history while listening to a musical piece, as well as on deliberate listening strategies. Brain structures underlying musical expectation are thus shaped by culture, as well as by personal listening history and musical training (Vuust & Frith 2008, p. 599).

This quotation not only indicates the relevance of expectation for the constitution of meaning, but involves most of the aspects that influence the process: subjectivity, formal musical structures, learning and culture, memory and temporal development. I will return to these aspects in the next chapters. The basic idea I try to support is that musical meaning is constituted in real-time processes and expectation is particularly relevant in determining such a meaning.

In its general sense, expectation is a tendency toward the future which is based on past experiences and learned knowledge. A musical meaning is the product of expectation when a musical event points to and makes us expect another musical event. In Meyer’s words, “the significance of a musical event – be it a tone, a motive, a phrase, or a section – lies in the fact that it leads the practiced listener to expect, consciously or unconsciously, the arrival of a subsequent event” (as cited in Levinson, 1997, p. 53). The Ideal Auditor’s expectations are based on the way she connects her knowledge of musical style with probability about future events based on statistical frequency. The actual expectations felt are thus a consequence of past experiences, which is followed by the constitution of evident meanings: “those which are attributed to the antecedent gesture when the consequent becomes a physico-psychic fact and when the relationship between the antecedent and the consequent is perceived” (Titchener & Broyles, 1973, p. 22).

In Meyer’s view, music is able to communicate meaning within a group when listeners share a set of gestures, knowledge and schemas. He argues that if a stimulus does
not arouse expectations, it is meaningless. Following this formalist point of view, the consequence is that music in a style with which we are totally unfamiliar is meaningless. In this case, listeners expect something, but they are not able to represent it correctly and to relate the events they are hearing to past meaningful experiences. Thus, the act of expectation which determines meaning during musical experience is culturally influenced.

Through tonal syntax, for example, listeners learn how to project forward in time. The tonal schema learned through simple exposure becomes fundamental in order to anticipate and understand music. As we have seen in other works (Lerdahl & Jackendoff, 1983a; Nussbaum, 2007), Western tonal art music since 1650 is based on a highly structured system of discrete pitches. Listeners learn how to organize tonal music according to its recurrent structures. Many empirical studies have demonstrated that listeners develop a sort of structural understanding at the perceptual level and that they represent music according to its syntax. Works like Haydn’s Symphony No. 94 in G major (Surprise Symphony, 1791) show how expectations and their violations influence the meaning the listener’s musical experience acquires. Due to the way the symphony is constructed, the work itself sets up some work-specific expectations that are then violated. For example, the main theme of the Symphony, which occurs in the second movement, contains an isolated fortissimo chord which does not commonly occur in slow, quiet music. An important factor contributing to the surprise is the fact that the composer has already presented this theme without the presence of the fortissimo chord.

\[ \text{\footnotesize I will develop this topic in detail in the fifth chapter.} \]
The sense of surprise is a cognitively ‘simple’ emotion, and a common response to unexpected stimuli which may be followed by other emotions such as fear, joy, or sadness. The emotional meaning the listener’s musical experience acquires is thus affected by expectation and its violation.

After an analysis of the most relevant theories about musical expectation, I will try to understand their role in cases where a listener is subject to unfamiliar music.
THE EXPECTATION PROCESS

Introduction

Music is composed using individual sounds which are then heard as a continuous connected whole. The listener is able to find relationships among the sound events occurring in the acoustical environment thanks to specific perceptual mechanisms, cognitive principles and neural processes. In particular, these allow her to integrate the sounds she hears into a structural whole and to understand the acoustical environment in terms of musical forms.

I discuss here the role of a specific mental mechanism, which seems to be very important to the listener’s musical experience and comprehension of music: expectation. As noted by Toiviainen and Krumhansl, “melodic expectations have been studied considerably because of the insights they offer to the basic perceptual processes involved in comprehending even the most elementary forms of music” (2002, p. 473). Taking into account the main theories about musical expectation that have been developed, I will focus on those aspects of expectation which highlight its role in basic musical understanding of the sound environment and in the constitution of musical experiences. To accomplish this task, I will consider the way in which the process of expectation works and how the creation of expectations influences the musical comprehension of an auditory event occurring in the acoustical environment.

During the process of organization that characterizes basic musical understanding, listeners create expectations and anticipations which determine their comprehension of music. This influences the way relationships between sounds are made. Concurrently, the dynamic nature of the process of expectation allows the listener to pick up on the developing organization of a musical piece.
I will also analyse several cognitive studies which investigate the way people typically hear pieces of music, and based on the results of these studies attempt to describe some perceptual aspects of the listening experience. Such experiments usually do not address the notion of expectation directly, but they are helpful in understanding how the expectation process may operate.\textsuperscript{60}

Expectation is a constant part of one’s mental life, and it is relevant to human psychology in all realms of experience. Studying how this process develops and determines the musical experience may reveal something more about the process of human knowledge in general. When listening to music or another familiar type of event occurring in the auditory environment, listeners experience a particular state of expectation which guides them to unconsciously anticipate future sound events. In its broader sense, expectation can be considered as a basic strategy of the human mind that reflects a tendency toward the future and is based on previous experiences. In music, it has often been related to the affective and emotional reactions a listener has during her musical experience. In this context I do not discuss whether music expresses, incites or contains emotions.\textsuperscript{61} If we

\textsuperscript{60} For example, Huron analyses some of the methods that are used in empirical research to investigate expectation. Through them we can understand the main aspects that characterize the expectation process and how it operates (Huron, 2006, chap. 4).

\textsuperscript{61} The relationship between music and emotion has been the subject of many philosophical debates concerning the aesthetic musical experience and the problem of music expressiveness (for an overview, Budd, 1985; Madell, 2002; Davies, 2001a). In particular, the central question concerns whether or not music can express emotion and how it might accomplish this. Different theories have been developed. The ‘expression theory’, for example, considers music’s expressiveness as depending on the composer’s intention to express emotions through the act of composition (see Davies, 2001: 32). The ‘arousal theory’ explains the expressiveness of music as its propensity to evoke the corresponding emotion in the listener by mirroring it (Matravers, 1998; Walton, 1988); the ‘new arousalist position’ claims that music arouses ‘garden-variety’ emotions whose intentional objects are musical events (Madell, 2002; Levinson, 1990). Kivy developed the ‘contour theory’ for which music expresses emotion by virtue of exhibiting similarities to human expressive behaviour (Kivy, 1989; Budd, 1995; Davies, 1994); in this view, expressiveness is a property of the music itself and harmony contributes to the contour of the music (Kivy, 1989).

Conventionalist theories relate the perception of musical emotions to learned musical conventions and arbitrary historical associations (Kivy, 1989), while others support the idea that the expressive power of music is grounded in certain features of the physical nature of sound, in particular the nature of the harmonic series (Cooke, 1959). The cognitivist view (also called ‘judgementalism’) represents the dominant analysis of emotion in analytic philosophy. Following Hanslick’s claim that music cannot ‘represent’ emotions since it lacks the conceptual means to do so, emotions are essentially considered as judgements thanks to which an indefinite state of mind passes into definite feeling (Hanslick, 1957; Lyons, 1980; Solomon, 1988).

Some authors regard music as a symbolic system which is expressive of something beyond itself: for this reason it is experienced as resembling or exemplifying what it denotes (Goodman, 1968; Langer, 1942;
consider ‘emotion’ as a kind of meaning that the musical experience can acquire, the study of how music induces emotions may be helpful in understanding how listeners build meaningful musical experiences. In order to investigate this last problem, I start the analysis with a brief historical review of the main theories about musical expectation that have been developed by music theorists and cognitive psychologists. I also consider the way the process of expectation has been studied through empirical research.

Main theories about musical expectation.

The role of expectation has been underlined by Leonard Meyer in his work Emotion and Meaning in Music (1956), where he seeks to establish and explain the general causes and conditions for affective aesthetic responses to music. Meyer begins by investigating the belief that an individual’s experience of music is derived from their own emotions and feelings about the music, which themselves are a function of relationships within the music itself. Therefore, he focuses his attention on the emotional response to music that is evident in the listener’s behaviour, and the ways which we may come to know more about the causal connection between the musical stimulus and the affective response it evokes.

In order to reply to this problem, Meyer proposes a psychological theory of emotions, which is a general hypothesis about the nature of the affective experience and the processes by which musical stimuli might arouse such experience. In this context, Sparshott, 1994). Langer, in particular, claims that music symbolizes the ‘general form of emotions’ and the principal artistic function of music is to symbolize feelings. From another point of view, musical emotions do not represent a real emotional experience because listeners perceive music as if it denotes or refers to an emotion that they have already experienced in the past. This is due also to the fact that music lacks a precise intentional object. In this way, the aroused feelings do not always correspond to the ones expressed by the music (Robinson, 1994). There is a level of uncertainty that allows subjective emotional experiences in listening to unfamiliar music as well, where the listener is more free in her interpretation and constitution of meaning.

I do not discuss these theories in detail, but follow the idea that music incites emotions because it moves, activates and stimulates in the listener particular cognitive mechanisms, such as expectation. The emotion felt by the listener is not constrained by arbitrary conventions and it is not a property of the music itself, but rather depends on her own experience- even in the case of unfamiliar music- and result in a subjective musical interpretation. From this point of view, musical sounds have dispositional properties which allow the listener to experience emotions on the basis of past emotional experiences as a result of specific mental processes.
expectations are important because they influence the affective response a listener has when she experiences a particular musical pattern.

Meyer argues that the emotional content of music arises from expectations and their ‘choreographing’ by the composer. This is clear in the Western tonal tradition by the way composers have used tonality to generate particular affective and emotional responses in listeners. In particular, he bases his theory on the thesis that emotions are aroused when a tendency to respond to a stimulus – an expectation – is arrested or inhibited. In Meyer’s words: “affect or emotion-felt is aroused when an expectation – a tendency to respond – activated by the musical stimulus situation, is temporarily inhibited or permanently blocked” (1956, p. 31). Music’s evocative power is thus derived from its capacity to generate, suspend, prolong or violate these expectations.

The ‘privileged role’ that musical expectancy has in the induction of emotions is also evident in other works. Juslin and Västfjäll, in their Emotional responses to Music: The Need to Consider Underlying Mechanisms (2008), consider musical expectancy as one of six mechanisms through which music listening may induce emotions. Musical expectancy “refers to a process whereby an emotion is induced in a listener because a specific feature of the music violates, delays, or confirms the listener’s expectations about the continuation of music” (2008, p. 568). Like Meyer, Juslin and Västfjäll support the idea that responses arising from musical expectancies also depend on sufficient exposure to the musical style in question. Auditory learning is therefore fundamental.

The creation of expectations is related to the knowledge a listener acquires during experience. In Meyer’s words, “[e]xpectation then is a product of the habit responses developed in connection with particular musical styles and of the modes of human perception, cognition, and response – the psychological laws of mental life” (1956, p. 30).

---

62 A tendency is a pattern reaction that operated in an automatic way. It consists of a set or series of regularly coincident mental or motor responses which follow a previously ordered course. In a broader sense all tendencies, even those which never reach the level of consciousness, are expectations.

63 Meyer distinguishes between “expectations that arise out of the nature of human mental processes – the modes in which the mind perceives, groups, and organizes the data presented by the senses – and those expectations that are based upon learning on the broadest sense of the term” (1956, p. 43). These two types of expectation are strictly related during the perceptual act. The former depend upon general cognitive and mental abilities, while the latter are culturally influenced.
This means that the listener’s expectations depend on the musical knowledge she has acquired during exposure to the musical style of her culture and during her past musical experiences in general. In this way, learning shapes one’s expectations over time. In Meyer’s view, listeners bring with them a vast body of musical experiences that, as one listens to a piece, conditions their response to that piece as it unfolds. Such a system of learned knowledge is not fixed but constantly altered by new experiences.

Another important theory about expectation has been developed by Eugene Narmour: the Implication-Realization (I-R) model of melodic expectations (1990; 1992). Narmour, like Meyer, analyses expectation in relation to the violation of regularities and he argues that expectations (musical and otherwise) stem from both innate and learned processing. His model is a detailed formalization based on Meyer’s work on expectation. The theory focuses on how implicative intervals set up expectations for certain realizations to follow. It is an alternative to Schenkerian analysis centered less on music analysis and more on cognitive aspects of expectation. In particular, in his book The Analysis and Cognition of Basic Melodic Structures (1990), Narmour formulates a comprehensive theory of melodic syntax to explain cognitive relations between melodic tones at their most basic level.

According to Narmour and Meyer, the structural aspect of music is particularly relevant for the creation of expectations. In a tonal context, for example, listeners develop a sense of musical expectation that is derived from tonal hierarchies (e.g. major and minor). This idea is followed by many other music theorists and cognitive psychologists. In Juslin and Västfjäll’s words, “musical expectancy refers to those expectancies that involve syntactical relationships between different parts of the musical structure” (2008, p. 568). Krumhansl and Agres (2008) also underline the influence of musical structure on the

---

64 Musical styles are more or less complex systems of sound relationships that are used and understood by a group of individuals. Based on such relationships the meaning of any term or series of terms depends upon its relationships with all other terms possible within the style system.

65 Schenkerian analysis is a method of musical analysis of tonal music based on the theories of Heinrich Schenker (1935). The goal of a Schenkerian analysis is to reveal the underlying structure of a tonal work (‘fundamental structure’ or Ursatz). The primary means of describing the structure of a musical passage for the Schenkerian analyst is to show hierarchical relationships among the pitches of the passage. This can be done through making reductions of the music and through a specialized symbolic form of musical notation that Schenker devised to demonstrate various prolongational techniques.
emotional response. They refer to Meyer, who focused his attention on the construction of the music itself. Vuust and Frith (2008) claim that the musical experience is dependent on the structures of the actual music as well, with additional emphasis on the expectations of the interpreting brain.

A further theory which supports this idea has been presented by Lerdahl and Jackendoff in their work *A Generative Theory of Tonal Music* (1983a). As we have seen in the first chapter, they try to explain which cognitive principles listeners use in order to organize the music they hear. To accomplish this they develop, in analytical form, an explicit musical grammar that reflects the musical knowledge listeners acquire during exposure to the music of their culture. This knowledge is unconsciously used to infer musical structures from the surface of all music to which they are exposed. In this context, expectation becomes an important mechanism for the comprehension of music.

Meyer uses the term ‘implication’ instead of ‘expectation’ when he writes: “I should now prefer to put these matters in somewhat more objective terms, referring to the implications which, given experienced and knowledgeable listeners, musical events have (or are felt to have) for one another, rather than expectations which listeners entertain about the future course of musical patterning. The difficulty with the term “expectation” is that it is often understood in a simplistic way” (as cited in Titchener & Broyles, 1973, p. 21). In this way he tries to give a more objective account of a process which is essentially subjective. I prefer the term ‘expectation’ because it reflects and better explains the subjective aspect of the musical experience.

Because of the uncertainty of musical communication and the probabilistic nature of musical style, Meyer relates the process of expectation to the information theory and the law of entropy. Peter Kivy explains this point of view when he writes: “information

66 This grammar, which takes the form of a system of rules, models the listener’s connection between the presented musical surface and the organization or organizations she attributes to the music. The kind of organization a listener attributes to a given piece, then, is not arbitrary but it is highly constrained in specific ways.

67 Considering those aspects of the musical structure that are hierarchical in nature and referring to Schenkerian analysis, the authors have identified four distinct hierarchical structures which are simultaneously imposed on a passage of music (the musical surface): grouping structure, metrical structure, time-span reduction and prolongational reduction.

68 See Meyer (1957).
theory says that events are on a continuum from the totally expected to the totally unexpected. The more expected an event, the less informative it is if it occurs, and vice versa: an unexpected event is ‘highly informed’, an expected event is not” (2002, p. 72). Juslin and Västfjäll, who consider musical expectancy in terms of brain function and syntactical processing, also make reference to the information theory. They regard the mechanisms through which music listening may induce emotions (including musical expectancy) as information-processing devices at various levels of the brain.

Donald Sherburne critiques Meyer’s point of view by noting that “Meyer’s theory seems to be incompatible with the ordinary conviction that fine music can be reheard and re-enjoyed many, many times, frequently with heightened appreciation” (1966, p. 580). For similar reasons Titchener and Broyles (1973) critique the information theory, referring to the ‘Information Theory Paradox’. The problem is that listeners are surprised not only when their expectations are not realized or satisfied, but also in listening to very well known music. Why does this happen? How are listeners surprised while listening to music that they have heard before? In order to reply to this fundamental question we have to take into account the fact that in every experience of a piece music there is something different: the context, the means of music transmission, and the emotive and mental state of the listener, to name a few examples.

In his book Sweet Anticipation. Music and the Psychology of Expectation, David Huron (2006) proposes a psychological theory of expectation, called the ITPRA (general theory of expectation), which attempts to explain how expectations evoke various feeling states, and why these evoked feelings might be biologically useful. He underlines the fact that the process of expectation depends upon auditory learning: listeners learn the regularities of the sound environment and they are sensitive to the probabilities of different sound events and patterns. These probabilities are then used to form expectations about the future. The violations of such expectations give rise to different kinds of emotions.

Expectations that are built on auditory learning reflect a statistical type of learning in which what a listener expects might simply reflect what she has experienced most frequently in the past. For example, listeners are faster at responding to scale degrees which
occur more frequently in the music of their culture. As Huron observes, “in a stable environment, the most frequently occurring events of the past are the most likely events to occur in the future. Thus, a simple yet optimum inductive strategy is to expect the most frequent past event. Acquiring such knowledge through exposure is referred to as statistical learning. The simple frequency of isolated events (“zeroth-order probability”) appears to provide the foundation for unconsciously learned expectations” (2006, p. 360).

In Huron’s opinion, expectations have a biological function and are culturally influenced. In the latter case, the regularities of the sound environment are learned in an unconscious way and are strictly dependent on the cultural context. From an evolutionary perspective, the capacity to form expectations and anticipate future events confers significant biological advantages. As the existing research shows, learning is important to evolution, and it occurs by changing the physical structure of the brain. Learning also influences the neural tracts of a single individual’s brain, which is fundamental to the dynamic creation of expectations.

Acquired knowledge and schemas allow accurate expectations about future events, and thus facilitate perception. It follows that the most expected event is the most pleasurable. As Annabel Cohen observes, “music appreciation relies on the mentally generated expectancies of notes soon to be heard, which music theorists and analysts have often tried to describe” (1982, p. 17). This observation is based on the idea that statistical regularities have a great influence on the way listeners understand and appreciate music. We will see that this is evident in the way listeners understand and appreciate music belonging to different and unfamiliar musical traditions.

Beginning with the necessity of having information about what individuals expect, Huron considers the methodological ways in which experimental studies are able to acquire information on the many factors which characterize a listener’s expectations. In most of

---

69 This is evident, for example, in the subjects’ speed of response (verbal or otherwise) to various experimental tasks.
70 The probabilistic character of expectation has been investigated in some studies by Temperley (2007, 2008); Deutsch, Henthorn, Marvin & Xu (2006); Creel, Newport & Aslin (2004); Saffran, Johnson, Aslin & Newport (1999). They present experimental evidence supporting the idea that listeners learn musical regularities and showing how probabilities shape music perception.
71 In this sense learning is not some disembodied social phenomenon, but an evolved neurological process in which aspects of the environment influence the microstructure of the brain.
these studies, expectation appears to shape many aspects of perceptual musical organization. They also suggest that expectations might have innate or learned origins that are often automatic and unconscious. I will consider some of these studies in the coming paragraphs.

Another theory in which expectation has a relevant role has been presented by Steve Larson (2002; 2004) in his theory of musical forces. He argues that listeners experience musical motion metaphorically based on their experience of physical motion in the environment. This metaphorical ‘hearing as’ is one source of expressive meaning in music which contributes to musical expectations. A related idea is that of ‘musical inertia’: a pattern of musical motion, once perceived as a pattern, will perceptually continue in the same way. Phenomena such as musical inertia reflect the fact that experienced listeners of tonal music (including jazz) learn to expect completions. Referring to Schenker, Larson underlines the role of musical structure and its hierarchical nature: “listeners expect music to complete the kinds of hierarchical structures that Schenker described” (2002, p. 363).

**Main features of musical expectation.**

When does expectation arise? What are its main features? How does it influence music comprehension? We can begin to understand how listeners build expectations by studying some of the mental mechanisms and processes used to organize the sound environment. Some of these are explained by the Gestalt psychologists; they are related to the nature of the human mind and thus considered innate for the most part. Other such mental mechanisms, however, are based on the way listeners learn the regularities of the music they are exposed to, and can be discovered through analysis of empirical research, In the latter case, for example, listeners use a set of ‘expectational schemas’ which are acquired during experience and used to categorize the music they hear. Musical context and certain specific musical cues (for example, social and environmental cues, stylistic cues)

---

72 As we have seen, in music schemas represent common enculturated aspects of musical organization.
are very important to this categorization process. With reference to the former consideration, some experiments have shown that an important Gestalt principle operating during auditory learning is the principle of grouping. As Meyer observes, the strength of expectation is related to perceptual grouping and it is entirely logical that the brain would group together stimuli that exhibit a strong statistical association.

In considering the main aspects of the expectation process, I will argue for its relevance to the constitution of musical experiences, and therefore on basic musical understanding. As we have seen, music comprehension is related to the ability to build mental schema and organize sounds into musical structures. According to Stephen McAdams and the studies which I referenced in the first chapter, during music listening, listeners store a hierarchical organization in their memory. When a particular musical structure (for example, a motive) is heard for the second time during a piece, the listener experiences a sense of development. In a piece where such thematic repetition occurs, the creation of expectancies is fundamental to the comprehension and appreciation of the music.

Individuals organize the music they listen to according to their knowledge, experiences and memory. If they are not able to do this, they cannot create expectations and the music does not develop. For some theorists it follows that listeners do not understand music the music in question. This idea could explain why listeners have a feeling of misunderstanding in listening to music they do not know or with which they are not familiar. If one experiences a piece of music which presents a complex organization or contains elements that they have never experienced before, it is very difficult to create relationships between sounds and to generate expectations. It is for this reason that the first time I went to a free jazz improvisation performance, I felt that I could not understand the music.

According to the theories previously discussed, some characteristic aspects of the mental processes involved in expectation are: auditory and statistical learning, the strong relation with the musical structure and the informative character of expectation. Expectations are also related to the listener’s intentions and beliefs and the sense of tension
and relaxation she experiences while hearing music. Tension, for example, is experienced as we approach expected events and prepares the subject for potential surprise.

Another interesting aspect of expectation is the sense of ambiguity (a state of mind in the listener) and uncertainty, which give rise to particularly strong tensions and powerful expectations.

In addition to the sense of tension, release, ambiguity and uncertainty which contribute to the definition of musical expectation, we must also consider aspects of attention. George Mead writes: “[o]ur whole intelligent process seems to lie in the attention which is selective of certain types of stimuli. Other stimuli which are bombarding the system are in some fashion shunted off. […] Our attention is an organizing process as well as a selective one” (as cited in Meyer, 1956, p. 83).

The ability to create expectations is related to the perceiver-environment relationship. This aspect emerges in Huron’s quotation: “it is possible to form relatively accurate expectations only because real-world events exhibit structure. The capacity to form expectations relies on the brain’s ability to create mental structures that emulate environmental structures” (2006, p. 358). Even if it is not clear what these environmental structures are for Huron, what is interesting in this quotation is that expectations are created thanks to relationships between the real-world, which contains its own structures, and the perceiver, whose perceptual structures have adapted to organize the information she receives from the environment into structural forms.

Memory, which is a fundamental requirement for all cognitive processes involved in music perception, plays an important role in the process of expectation: listeners generate expectations about the future on the basis of the knowledge that they retain in memory and recall when particular events occur. Huron, for example, discusses the link between expectation and memory by analyzing three different forms of expectation in relation to three types of memory: episodic memory, semantic memory and short-term memory.73

73 ‘Veridical expectations’, for example, are based on long-term patterns arising from repeated exposure to a single episode, token, or work. ‘Dynamic expectations’ represent short-term patterns that are updated in real time, especially during exposure to a novel auditory experience such as hearing a musical work for the first time (2006, chap. 12).
As already underlined, the process of learning is particularly relevant for the creation of expectation, which in turn is developed due to the essential contribution of memory. Levinson writes: “before the first note is heard, then, we have a listener who has certain expectations, though not necessarily conscious ones, as to what he will hear” (1997, p. 45). He notes that these expectations derive principally from two sources: past experiences of a particular musical style and past experiences of a specific piece of music. The listener’s reaction to a bit (a present perceived passage of music) is affected by her memories of past experiences and the previously heard bits, and by the norms of tonal music which she has unconsciously acquired during exposition and through repetition.

As Levinson observes, when listening to a piece of music “the listener’s expectations for the third phrase will be a product of at least four things: (1) internalized knowledge of how pieces of a certain sort are likely to continue, that is, of stylistic probabilities; (2) innate psychological tendencies to regard certain continuations as more likely than others; (3) stored representation of the remainder of the piece from earlier auditions; (4) memories, more or less vivid, of what has been heard earlier in the piece on the present occasion” (Levinson, 1997, p. 48). This quotation not only summarizes some of the most relevant features of the expectation process, but it also introduces a fundamental requirement for the creation of expectations during musical listening: the temporal structure of human consciousness and the temporal development of the listener’s musical experience. I will analyse this in detail in the sixth chapter.

Expectation also has a strong subjective character which determines the way listeners develop musical experiences from sound events occurring in the acoustical environment. We have seen that, according to Huron, expectations depend on statistical learning and on the degree of probability that something will happen. During their own experience, listeners build a kind of hierarchy of possibilities that determine a subjective creation of expectations. As a consequence, the degree of satisfaction with the outcome of these expectations influences the comprehension of music. For example, if my expectation of hearing a cadence at the end of the second movement in a sonata form – an expectation
which I have built on the basis of learned voice-leading\textsuperscript{74} tendencies in Western tonal music – is satisfied, I will have a sense of comprehension and familiarity with that musical event. The probability that my expectations will be satisfied is related to my previous knowledge. Expectations and their satisfaction thus depend on a subjective structural organization of sounds. It is important to impress that expectations depend on past knowledge, but they are distinct from knowledge about the future because the future is not yet experienced, nor perceived, thus only probable. As Titchener and Broyles observe

none of these writers [Meyer, for example] distinguishes adequately between the experienced anticipation that the Ideal Auditor feels as he listens to one of the inventions and his knowledge that the invention will end in a precise way. Once the Ideal Auditor knows a piece well, he knows how it will end and that the ending will approximate a stylistically accepted pattern. But all this knowing is different from the actual experiencing of the ending itself. It is conceivable, for instance, that the Ideal Auditor might know things of this sort without having experienced them directly. He has, let us suppose, never heard the F-major invention, although he knows how it is likely to end because he has heard and studied the other inventions. Here the distinction is drawn between knowing an ending and experiencing it (1973, p. 18).

In this way, Titchener and Broyle make a distinction between knowing an ending and experiencing it, underlining the unconscious aspect of the expectation process and its ability to make reference to what is not yet present. This also introduces a distinction that I will develop in detail: the difference between the expectation of a potential future event and the anticipation of a known future event.

Two questions related to these observations are: are listeners always able to form expectations? What happens when a listener experiences unfamiliar music? A difficulty in addressing such questions on the basis of the available literature is that most of the theories and empirical studies focus only on tonal music and do not give good insights into the study of how atonal music or music that is not well known are understood. Another problem is that it is not clear how expectations determine comprehension. It seems that listeners always create expectations, but if a listener is not able to create ‘correct’ expectations, music becomes meaningless and she has no comprehension. I do not agree with the notion that it is necessary to have ‘correct’ expectations in order to understand

\textsuperscript{74} Voice leading refers to the way individual notes typically move from chord to chord.
music (I will discuss this point in the fifth chapter), however this perspective does follow
the idea that when expectations are not satisfied, listeners do not understand or are not able
to have meaningful musical experiences.

**Empirical studies about expectation.**

In order to understand the main features of the expectation process and the way it is
developed, I will take into account some of the empirical studies that have investigated its
role in comprehension and emotional experience. Following Meyer’s theory, some
researchers have investigated in an empirical manner the ways in which musical expectancy
may induce emotions or evoke various feeling states.75 Most studies currently available,
both in the field of music theory and the psychology of music, have analysed musical
expectation solely in relation to the melodic76 and rhythmic dimension of music,77 without
paying much attention to other musical dimensions (such as timbre).

By studying the *when* and *what* of something which is expected to occur during the
unfolding of a musical event, empirical studies have investigated the ability of a listener to
perceive and identify melodic and rhythmic patterns, as well as the influence that
expectations and their violations have on this ability. In doing this, they refer almost
exclusively to Western tonal music and analyse the expectations that a listener creates on
the basis of the musical structure which the composer has imposed on the piece of music.
The creation of patterns of tension and relaxation, for example, results in a structural
organization of sounds that brings about particular expectations.78 One problem involves
the question of how these studies show and highlight the features of the expectation process
that we discussed in the previous paragraph.

---

75 See, for example, Juslin & Västfjäll (2008); Longhi (2008); Krumhansl & Agres (2008); Vuust & Frith
(2008); Sloboda (2002); Gabrielsson & Lindström (2001).
76 See, for example, Unyk & Carlsen (1987); Larson (1993; 2004); Lerdahl (2001); Margulis & Levine
(2004); Margulis (2003); Aarden (2003); Krumhansl (1979).
78 See Eerola, Toiviainen & Krumhansl (2002); Margulis (2005); Margulis & Levine (2004).
In one of their studies, Eerola, Toiviainen and Krumhansl (2002) asked participants to rate the predictability of melodies while the melody was playing. More precisely, they asked participants to rate the probability of occurrence of particular events in the melody – such as pitch-classes, intervals or two-tone transitions – using a slider on a computer screen. Listeners had to indicate the more predictable events at each moment by moving a mouse-driven slider on the computer screen: the far left end of the slider was labelled as very predictable and the far right as very unpredictable. Through a dynamic approach and the use of three models for melodic predictability, the authors were able to investigate common melodic expectations by analyzing participants’ errors, which revealed what they actually expected at each point in the melody. The authors conducted their experiments with musically experienced participants only (25 music students from the University of Jyväskylä). Because of their musical background, these subjects were trained in listening to tonal music and had specific knowledge about tonal structures. The analysis of the results (one of which is presented in the figure below) indicates that learning and musical knowledge are highly influential toward the creation of musical expectations and that expectations are dependent on the musical material.

---

79 As the authors writes, “model 1 consisted of Narmour’s implication-realization principles supplemented with tonal stability and melodic anchoring. Model 2 was based on recursive probability distributions of two-tone continuations where the initial probabilities were derived from a large corpus of folk melodies. Model 3 comprised the entropy of the distribution of the tones within a moving window” (Eerola, Toiviainen & Krumhansl, 2002, p. 473).
Another work providing evidence that certain rules of melodic expectation are learned rather than innate has been presented by Paul von Hippel (2002). He compared the expectations of musicians and non-musicians in relation to tonal rules concerning structural elements of a musical composition. An example of a learned rule is the expectation for common rather than unusual scale degrees – for example, diatonic rather than chromatic tones. During the experiment, participants had to hear a melodic fragment and answer the following question: ‘if this melody were to continue, would it be more likely to go up or down?’ The results of this experiment showed that while the musicians’ expectations corresponded to well-known voice-leading rules, such knowledge was not evident in the expectations of non-musicians. This suggests that prediction strategies are learned through
engagement with music. An interesting point made by Hippel is that learned strategies reflect imperfect prediction strategies (heuristics). In the paper *The Influence of Expectancy on Melodic Perception* by Anna Unyk and James Carlsen (1987), the authors review some experimental research about musical expectancy and present a study they conducted with the purpose of further testing the relationship between expectation and perception within the context of a melodic dictation task. They observe that “through experience with styles, listeners internalize these probabilities of occurrence which become the basis of their expectancies for future events while listening to unfolding music patterns” (Unyk & Carlsen, 1987, p. 3). This shows that listeners unconsciously learn schema and patterns that they use to form categories, organize what they hear and abstract the underlying structures of music. They also recall a study by Carlsen (1981) where subjects were asked to sing what they imagined the continuation of an interrupted melody would be. Carlsen found that different expectancies were sung in relation to the proposed melodic beginnings. He also noticed that the sung continuations were different for subjects coming from three different countries. The results support the idea that music expectancies are dependent on previous experiences with music. Moreover, Unyk and Carlsen write: “if expectancy is defined as the anticipation of future events in an unfolding music pattern, and the content of these expectancies are found to differ across subjects, then a measure of the effects of violation of expectancy on perception must also reflect these individual differences” (1987, p. 14). This assessment indicates their attention to the subjective character of musical expectations.

Elena Longhi (2008) has conducted experiments focusing on the structural elements of music in relation to musical expectancy. She notices that when mothers sing to their 3-month-old infants, they emphasize some aspects of the hierarchical structure of the songs, thus creating expectations imbued with emotional power. For example, a detailed analysis revealed that the mothers emphasized the metrical and phrasing structure of the songs. Extending the duration of certain beats over others, they marked the boundaries between

---

80 In psychology, heuristics are simple, efficient rules which can be hard-coded by evolutionary processes or learned. They have been proposed to explain how people make decisions, come to judgments, and solve problems, typically when facing complex problems or problems with incomplete information. These rules work well under most circumstances, but in certain cases lead to systematic errors or cognitive biases.
Another study which investigates melodic expectation is Margulis’ *A Model of Melodic Expectation* (2005). Her model assigns rating to the expectedness of melodic events, which depend on the hierarchic implementation of three primary factors – stability, proximity, and direction – and one secondary factor – mobility. Expectations are linked to the listener’s experience of tension in the melody. The article assesses the role of expectation within the cognitive processes that underlie perception and suggests methods for empirically investigating expectations and the real-time experience of listening. She observes that “music does not seem merely a linear succession of more and less tense junctures; rather, it seems qualitatively rich and multidimensional. Expectancy-tension is an inherently forward-looking, prospective phenomenon, and events with high values for it should seem saliently implicative” (2005, p. 697).

Hazan and collaborators (2007) present a system and a computational model that produces expectations for rhythmic music signals at a constant tempo. The system is able to learn the hierarchical structures of the presented musical sequences and, on that basis, it builds expectations.

The authors aim to obtain a system which can be used in the context of modelling musical expectation and memory from both computational and cognitive points of view. In the work *What/when Causal Expectation Modelling in Monophonic Pitched and Percussive Audion* (2007c), Hazan et al. present a causal system for representing a musical stream and generating further expected events. They focus on the musical dimensions of timbre and rhythm. In particular, they “propose a causal and unsupervised system that learns the structure of an audio stream and predicts its continuation” (2007c, p.1).

---

81 Margulis’ model further extends Narmour’s Implication-Realization model and describes three distinct types of listener reactions, each derived from listener-experienced tension: ‘surprise-tension’, which is inversely proportional to degree of expectancy and results in intensity or dynamism; ‘denial-tension’, which is proportional to the discrepancy between the expectancy of the most expected event and the expectancy of the actually perceived event, and results in desire, drive, will; ‘expectancy-tension’, which is proportional to the degree of expectancy of the most expected event (in other words, if the listener had no idea what to expect next, the expectancy-tension would be low), and results in strain or yearning.
In their paper *The Influence of Expectancy on Harmonic Perception* (1992), Anderson and Tunks investigated the relationship between music expectancy and harmonic perception, as measured by a harmonic dictation task. They tested the hypothesis that listeners’ perception accuracy would be greater for chords fulfilling their expectations than for chords violating their expectations. The results of their experiments indicated that the greatest number of errors in harmonic dictation did indeed occur on low-expectancy chords. The fewest errors occurred on high-expectancy chords, and an intermediate number of errors occurred on moderate-expectancy chords. This indicates the ability to understand and reply in an appropriate way to musical events which are more familiar. As Anderson and Tunks conclude: “[r]esults of this study add support to existing theories of music expectancy stressing the importance of expectancy in music perception and cognition. What people expect to hear seems to have an influence on what they do hear” (1992, p. 11). They were also able to find significant differences in expectancies between subjects as a function of their musical training and knowledge. Different expectations are thus related to different degrees of exposure and musical training.

David Hargreaves (1986) has proposed the concept of ‘subjective complexity’ to indicate a relative scale of composite music complexity which is based on an individual’s personal musical experience. The basic idea is that music which is subjectively simple to an experienced musician could be subjectively difficult to a musically naïve listener. Following the information theory, music which is too subjectively simple contains little surprise and is not musically appealing. Alternately, music which is too subjectively complex contains too many surprises and does not make sense to the listener.

These studies, among many others, highlight some of the features that characterize the process of expectation and its relevance to musical listening. These features include: auditory and statistical learning; the relevance of musical structure in the creation of expectation; the influence of cultural exposure; the informative character of expectation; the role of tension and relaxation in arising expectation; memory and attention as influential mental mechanisms; intentions and beliefs; the temporal development of the musical experience and the subjective nature of musical experience. These factors are all very
important to the constitution of musical experiences and to basic musical understanding at the perceptual level.

**Some reflections.**

Most of the theories and studies about the process of expectation outlined in this chapter try to give a causal account of the way listeners create musical expectations. Meyer, for example, looks for causal connections to support his hypothesis about the influence of expectation on affective musical experience. He also suggests an objective discussion of a subjective content (emotion) through the examination of the music itself.

Based on the topics covered in the first and second chapters, my question is: is it possible to give a causal account of a subjective experience and the processes through which it is built? Meyer himself, in his successive works, substitutes the objective terminology he uses in *Emotion and Meaning in Music* with a more subjective one. He makes this change based on the realization that it is necessary to speak in terms of expectations generated by a properly attuned listener, or in terms of the probabilities a properly attuned listener unconsciously makes about what musical events will follow those he is hearing.

Hanslik claims that “there is no causal nexus between a musical composition and the feelings it may excite, as the latter vary with our experience and impressibility” (as cited in Meyer, 1956, p. 270). We can easily notice that listeners have different affective experiences in relation to the same piece of music: some of them have a positive emotional response, while others might have a negative response. Why does this happen? How can we account for this difference? Is it always possible to find causal relationships that explain the subjective experience and thus anticipate a listener’s emotional behaviour?

Considering that emotional experience depends on the realm of one’s mental activity, the differences in subjective experiences lie in the relationship between the stimulus and the responding individual. Listeners may use very different mental codes to
perform the same task, and this could determine different individual responses. In other words, listeners have different subjective musical experiences because of the way they process sound events. For this reason it is not always possible to explain such experiences in a causal way.

From a phenomenological point of view – which takes into account the overall and subjective perceptual experience – the role of subjectivity is highlighted by the concept of constitution of musical experiences, which I discussed in the second chapter. Most of the theories that I have taken into account relate expectation to the emotion a listener has while hearing music, which might be considered as the meaning her musical experience acquires. One of the purposes of the present work is to explain the influence (if any) that the process of expectation has on the listener’s musical comprehension of the sound environment and in the constitution of musical experiences.

As we have seen, music comprehension involves different kinds of musical understandings and musical experiences. For example, it could be related to the meaning a piece of music has within a specific culture, to the understanding of specific musical passages and structures, or to the emotive character a piece expresses for an individual. At its basic level, comprehension is related to the capacity of a listener to have musical experiences and organize the acoustical environment into musical forms. I will argue for the idea that expectation influences such comprehension and constitution of meaning in different ways.

An interesting feature of the expectation process is that it is an *intentional* movement oriented toward the future. As we have seen, intentionality characterizes an important aspect of basic musical understanding and the constitution of musical experiences. Expectation appears to be a function of the human mind-brain that reflects an intentional movement toward the future – a teleological dimension which has relevant consequences for survival and human knowledge about the world. This definition could be related to a teleological perspective for which expectation is a biological function which allows for anticipation of future dangerous events. In this sense it is very important for survival.
As I have previously highlighted, theorists give musical structure a great role in the creation of expectations. In their opinion, the ability to find structural and hierarchical relationships among sounds seems to be fundamental in order to have expectations about future events. Such relationships depend on the compositional structure of the music itself. From this perspective, listeners are able to create ‘correct’ expectations if they have already had experience with the music they are listening to. For example, it is easier for them to recognize and have musical expectancies about the music of the culture to which they have been exposed. Emotional experience is also related to one’s knowledge of a particular musical idiom, in that emotions arise due to violations of expectations. According to this view, one’s ability to have musical comprehension depends on correct understanding of the musical structure.

In my opinion, a limit to this approach is that it relies on music theory and analysis, which assumes that the perceived musical structure is the same as the notated one. It follows that if the listener is not able to perceive the notated structure, she cannot understand the music. However, a non-musician – that is, one who does not have conscious knowledge of music theory – is often able to grasp structures which are typical of her culture’s music because of learned musical knowledge. But if she does not understand the structural relationships adequately (as defined by the theorist’s perspective), does it mean that she is not able to create expectations, and, therefore have musical comprehension?

When listeners experience music belonging to a different culture, they have no specific knowledge (implicit or otherwise) about its structure. They organize it on the basis of their previous experiences. In organizing the unfamiliar music in this way, it might happen that they do not understand the correct musical structure, but may still have an auditory experience which is musically meaningful. Huron presents an example of missing comprehension as he reflects upon an instance of misinterpreted musical meaning. Based upon his Western tonal exposure-based notion of musical structure, Huron misconstrued the musical genre of a Lakota Native American song as being a lullaby when it was actually horse-stealing song. This error occurred because he did not know the language or the rules used to create the musical structure of the song in the tradition of Lakota musical culture.
However, he still experienced the music as having a specific meaning (that of being a lullaby) as a result of the expectations he had been able to create based on his own knowledge and schema.\textsuperscript{82}

What happens with completely unfamiliar music, like atonal music for an ‘expert tonal listener’? As we will see in detail, in this case listeners trained with tonal music continue to hear the musical sequences according to tonal expectations. This has some implications for the comprehension of music and the meaning their musical experience acquires. In particular, the perceived structure does not necessarily correspond to the notated one. This firstly shows that the musical experience is a subjective matter (which implies different kinds of comprehension), and secondly, that it is not possible to give a completely causal account of such a subjective experience.

Another issue that I would like to take into account concerns a difference between Huron and Meyer’s theories in relation to the problem of unfamiliar music and its appreciation. Meyer claims that emotions arise when expectations are broken and, in the case of unexpected events, listener’s acquire more information. If, on the other hand, what is expected does happen, listeners do not have new information and, from Meyer’s point of view, they do not experience new emotions. If this were true, however, listeners would not feel emotions in listening to very well known pieces of music.

Huron seems to follow the same line of thought as Meyer, but from a biological point of view, a broken expectation represents a failure. He also observes that listeners enjoy music when their expectations are fulfilled. A related problem concerns the reason why a person may listen to the same piece of music many times and still have emotional experiences, whether they be the same or different. We need to make a distinction here between emotional experience and the enjoyment of music in order to better understand the role of expectation and what happens in the case of unfamiliar music. If a listener has an emotional response while listening to unfamiliar music or unexpected events, it does not mean that she necessarily enjoys the music. On the other hand, she could enjoy music even if she does not experience any kind of emotion. According to the former case, for example,

\textsuperscript{82} See Huron, 2006, p. 215-216.
listeners can experience strong emotions with unfamiliar music (even if they do not appreciate it) because their expectations are broken. This case becomes problematic if considered within Meyer’s theory: if emotions arise when an expected musical passage in a particular musical style is not fulfilled, then in the case of unfamiliar music listeners would theoretically not be capable of experiencing emotions as a result of their inability to create expectations.

Based on these considerations, I would like to better analyse the role of expectation in listening to unfamiliar music and which kind of musical experience it determines. For this reason, in the next chapter I will introduce a conceptual distinction between ‘expectation’ and ‘anticipation’.
CONCEPTUAL DISTINCTION:
EXPECTATION VERSUS ANTICIPATION

Introduction

The majority of the theories and empirical studies about musical expectations that I have outlined thus far use the words expectation and anticipation almost synonymously. For example, Unyk and Carlsen claim that “the process of expectation or anticipation of future events in listening to music is the focus of several theories of music perception” (1987, P. 3). Authors like Vuust and Frith do not make any distinction at all, as we can observe in this sentence: “if we consider music expectation/anticipation as the fundamental mechanism for the musical experience ….” (2008, p. 599). Underlying the dynamic aspect of expectation, Ralph Barnes and Mari Riess Jones claim that “at a general level, we think that a dynamic attending approach to expectancy conforms to popular intuitions about the way we relate to many changes in our experienced environment. This is because expectancy typically implies a future-oriented aspect of one’s attentional set. […] And we expect the next note in a melody to sound a few hundreds of milliseconds in the future. Regardless of differences in time scale, expectancies are conditional on time: they reflect anticipations about something that is destined to occur within some (broad or narrow) temporal neighbourhood in the future” (2000, p. 302). In my opinion, it is not clear how any of these authors distinguish between expectation and anticipation and to what anticipation refers. Is it a kind of perception, a mental representation or an act of imagination?

In this chapter I discuss the possibility of a conceptual distinction between expectation and anticipation, and the idea that such a difference could give new insights for a better comprehension of the subjective musical experience. It will also be helpful to clarify what expectations are and define their role in the basic musical understanding of
sound events. In particular, I propose that such a distinction might reveal interesting aspects of particular musical experiences, such as listening to unfamiliar musical events.

**Discussion**

We could start the discussion by making a terminological distinction between the two words. Etymologically, ‘expectation’ is a French word which has its origin in the Latin word ‘expectation’ (expectation, suspense) and ‘expectare’ (lookout for, await, expect). It indicates a state of suspense, a sense about and a tendency toward the future. ‘Anticipation’ originates from the Latin word ‘antecapere’, which means ‘to take, to grasp before’ (ante = before; capere = to take, to grasp). It may be understood as the action of representing an event before its occurrence. Thus, anticipation seems to be more related to an action, while expectation to the belief that something will happen. In Western tonal music theory, an ‘anticipation’ refers to the introduction of a note before its expected place in a harmony (i.e. the action of playing a note before its correct occurrence in time). If we consider the listener’s point of view and her perceptual experience of a musical event, the etymological distinction is not satisfying as it does not explain the influence of expectation on the constitution of musical experiences. For this reason, I propose an in-depth analysis of a few studies on the psychology of expectation.

In the Glossary of his book *Sweet Anticipation: Music and the Psychology of Expectation*, Huron gives the following definition of ‘anticipation’: “1. The subjective experience accompanying a strong expectation that a particular event will occur; also referred to as the feeling of anticipation. 2. In Western music theory, a type of melodic embellishment in which an expected note is immediately preceded by the same pitch. E.g. The “ta” in the “ta-dah” cadence” (2006, p. 409). He also relates ‘anticipation’ to ‘premonition’, which he defines as “a long-range feeling of anticipation” (2006, p. 418). In the Glossary we do not find a definition of ‘expectation’, even though both expectation and anticipation are fundamental concepts in Huron’s theory. They both appear in the title of
the work (*Sweet Anticipation. Music and the Psychology of Expectation*) in a way that suggests Huron views anticipation as a particular kind (sweet anticipation) or part of a general process of expectation (for which there is a ‘psychology of’). Through a careful analysis of the book, we can see that ‘sweet anticipation’ has a specific technical definition: it refers to the positive feelings that arise from conscious thought about some future events, such as the thought of attending a concert. It is considered as a special subset of expectation. Unfortunately, very little of the book deals with particular types of expectation, and the subject matter of the work is better identified by the subtitle *Music and the Psychology of Expectation*. However the very presence of a definition of ‘anticipation’ means that Huron makes a distinction between the two terms in question, even though they are not clearly distinguished in the work. The problem left to understand is what kind of distinction he makes and why this distinction may be relevant to a better comprehension of the musical experience.

Huron’s definition of anticipation underlines the subjective character of the musical experience. This is in line with the idea that expectations are created on the basis of one’s own way of listening to music and her cultural influences – an aspect of the expectation process that we have seen in the previous chapter and which is highlighted by both theories and empirical studies about musical expectancy. Another interesting aspect emerging from the definition of ‘anticipation’ is the reference to the object of the anticipation, which seems to be a particular – and therefore specific – event.

As previously mentioned, according to Western music theory, anticipation is the introduction of a precise note before the expected moment. In the case of musical practice, for example, anticipation occurs when a note is played before the chord to which the note belongs and resolves when the ‘anticipated’ chord is reached:

![Figure 3](example_of_anticipated_note.png)  
*Figure 3. Example of anticipated note.*
Focusing on music perception, an important question arises: is it possible that the anticipation of a note might correspond to the act of mentally representing it – consciously or unconsciously – before its correct time of occurrence and immediately after a strong expectation that the note will occur? In this case, is it still possible to speak of a general ‘feeling of anticipation’, or is it better to refer to it as a specific act of consciousness? The basic idea is that anticipation is a kind of mental representation that follows a strong expectation about an impeding event and determines the subjective musical experience. As we will see, the listener does not always anticipate, as the creation of anticipations is related to various factors. The problem remains to understand what kind of act this is and if it is experienced at a conscious or unconscious level. I propose to develop such a problem by clarifying the aspects which differentiate anticipation and expectation, and by discussing how these influence the constitution of subjective musical experiences. By consequence, I will outline how both anticipation and expectation affect one’s basic musical understanding of auditory events occurring in a complex acoustic environment.

I will better analyse the distinction between expectation and anticipation with respect to their possible influence on music perception in the sixth chapter, but for the moment I will only introduce the problem through an account of David Huron’s work.

Trying to understand what ‘having an expectation’ means, Huron addresses the questions: “how precise are expectations? Do we expect specific events, or do we expect “classes” or types of events?” (2006, p. 41). He firstly replies to such questions by observing that “one definition of expectation might classify it as a form of mental or corporeal “belief” that some event or class of events is likely to happen in the future” (2006, p. 41). The term ‘belief’ may refer to a range of expected possibilities with different degrees of certainty about the occurrence of a future event. Beliefs about the future are evident in a person’s ‘action-readiness’ – that is, changes of posture, metabolism, or conscious thought that prepares the individual for certain possible outcomes but not for others. Expectation, then, allows a listener to believe that something will happen and to prepare herself to respond, but this does not mean that she represents a specific event in her
mind. It refers to a series of possibilities and it indicates a mental state in which the most likely events to occur are expected.

I suggest that the anticipation of an event is distinct from the belief and feeling that it will occur. What is different is the specificity of the object, how it is represented before its occurrence and the way the subject prepares herself for its outcome. I will return on this important distinction in the sixth chapter.

In the fifteenth chapter of *Sweet Anticipation. Music and the Psychology of Expectation*, Huron focuses on the feelings that precede highly expected events. To explain and characterize these feelings, several researchers have proposed specific terminology. For example, Bharucha has suggested the term ‘yearning’; Narmour uses the term ‘implicative’; and Margulis the technical term ‘tensions’. Huron uses the phrase ‘feeling of anticipation’ in order to distinguish the sense from the Western tonal music phenomenon previously outlined. He writes: “I will use the phrase “feeling of anticipation” to refer to the sense of expectancy. When a listener is certain of some future event, we may say that the listener is experiencing a strong feeling of anticipation” (Huron, 2006, p. 306). Anticipation is thus seen by Huron as a special subset of expectation in which there are some impending events that are entirely expected. There is a strong feeling that accompanies such expected certainty, which he has called the ‘feeling of anticipation’.

Huron’s definition suggests that anticipation can be related to some specific events and the ‘act’ of anticipating them (for example, through a mental representation or a specific ‘action-readiness’), while expectation is a ‘sense’ about a future which is not clearly defined. “In many ways, expectation can be regarded as yet another sense: a sense of future” (Huron, 2006, p. 355): it provides the mind with information about upcoming events. My question concerns whether such information is followed by some sort of mental representation that precedes the expected event. Could we call this mental representation ‘anticipation’?

Looking again at Huron’s ITRPA theory of expectation, he distinguishes five expectation-related response systems that are evoked at different times during the expectation cycle; these are divided into pre-outcome and post-outcome responses with
respect to the event onset. The pre-outcome responses are: the imagination response and the tension response. According to Huron, the former has the purpose of motivating an organism “to behave in ways that increase the likelihood of future beneficial outcomes” (2006, p. 15). In imagining different outcomes, feeling states are thus activated. The latter has the purpose of preparing an organism “for an impending event by tailoring arousal and attention to match the level of uncertainty and importance of an impending outcome” (2006, p. 15). As an expected event approaches, physiological arousal typically increases, often leading to a feeling of increased tension.

The post-outcome responses include: the prediction response, which provides positive and negative inducements that encourage the formation of accurate expectations and evokes feelings in relation to whether one’s predictions were born out; the reaction response, which addresses a possible worst-case situation by generating an immediate protective response; and the appraisal response, which provides positive and negative reinforcements related to the biological value of various final states.

![Figure 4](image-url)  
**Figure 4.** The schematic time course of Huron’s five expectancy processes (Scott, Tsou, Schmuckler & Brown, 2008, p. 138).
As is clear from the ITPRA schema, in Huron’s view expectation is a complex cognitive process which is characterized by very specific moments operating in a continuous chain of dependent responses preceding and following the occurrence of a particular event. I focus my attention only on the pre-outcome responses in order to define ‘expectation’ in its narrowest sense. I argue that the definitional difference between ‘expectation’ and ‘anticipation’ may influence the way we think about how music is understood at the perceptual level. In particular, I propose to consider both expectation and anticipation as imagination responses which maintain different functions. The anticipation imagination response, for example, is very important to survival: in a dangerous situation it is more important to anticipate a specific event than to expect a range of possible events, as one would with the expectation imagination response. Because of the different degrees of strength through which expectation and anticipation refer to the event onset, they influence the tension response. This is clear if we observe the listener’s ‘action-readiness’ and the way she prepares herself to react to the event (for example, a listener shows a different bodily gesture when she anticipates a musical event than when she expects it). By consequence, they determine different post-outcome responses.

In the case of expectation, I, as a music listener, feel that something has to happen and so I prepare myself to react to it, but I am not sure what the ‘something’ is. In relation to anticipation, on the other hand, I prepare myself for specific events. From this point of view, expectation is related to a mental state, while anticipation determines a precise mental representation about what will happen. In this sense, anticipation is also related to a kind of ‘mental action’ or perception which is similar to the actual perception of early sounding ‘anticipation’ tones as defined by music theorists.

From a cognitive point of view, expectation is a mental process that implies some sort of mental representations – auditory images\(^{83}\) – of the sounds that are analysed,

\(^{83}\) An auditory image is defined by McAdams in this way: “L’image auditive est une représentation psychologique d’une entité sonore révélant une cohérence dans son comportement acoustique” (1985, p. 2). Moreover, “imaging proceeds as a kind of collecting of things that “belong together”. The image emerges as a representation of this collection of the parts of a sounding body distributed across time and frequency” (McAdams, 1982, p. 280). The creation of auditory images evoked by the acoustic environment is an
interpreted or distilled.

Do these images refer to specific musical events? In this case, could we say that expectation corresponds to a mental state – a sense about the future – which is followed by a kind of representation of *what, when* or *where* the musical event is expected to occur? I propose to consider such representation as anticipations.

From a temporal point of view, even if expectation and anticipation are both part of the imagination response, anticipation always follows a strong expectation. This means that a listener may expect some future events without anticipating a precise one, but not the opposite. As we will see shortly, this difference may be one of the reasons listeners have a feeling of misunderstanding when they encounter unfamiliar music.

In the field of music cognition, melodic expectation is generally defined as the tendency to have a feeling about what might come next in a melody or succession of harmonies. For example, if the ascending musical partial octave ‘do-re-me-fa-sol-la-ti...’ is heard, listeners familiar with Western tonal music will have a strong expectation to hear one more note, ‘do’, in order to complete the octave. I propose to consider the general feeling of moving toward a musical goal as an expectation, and the strong expectancy to hear the specific note ‘do’ as an anticipation.

One method used to study the difference between musical anticipation and expectation is through the analysis of the way a listener sings along with the music she is hearing. In a case where she is familiar with the music, she uses anticipation when she sings before a melodic note is actually heard. In this case, could anticipation be thought of as an act (singing before the correct time) that corresponds to a specific mental representation? How is this mental representation characterized?

Expectations, in Meyer’s view, are not specific, but rather indicate a general state of suspense: “what is expected in this state of suspense may not be specified, but this does not mean that any consequent is possible” (Meyer, 1956, p. 29). Listeners sense that something will happen, and although they are not certain of exactly what the expected

---

84 Meyer considers suspense essentially as a product of ignorance about the future course of events. It arouses strong mental tendencies toward clarification which are immediately affective. (See Meyer, 1956, p. 27)
musical event will be, they are sensitive to a range of possibilities which are built on the basis of previous experiences. This may be the case because, at a certain moment in the musical experience, they are not able to have precise mental representations. Suspense, in fact, is a product of ignorance about upcoming events, from which strong mental tendencies toward the future arise. Expectations themselves do not coincide with specific mental representation of future events.

Therefore, I propose to consider expectation as a mental state through which listeners prepare themselves for future musical events, and anticipation as the action of representing one of them before its occurrence. The entire process – that is, expectation followed by anticipation – is built on the basis of learned knowledge and previous experiences. In the case of expectation, listeners expect some events on the basis of their probability of occurrence. In anticipation they represent in their mind a specific event if the expectation they have had was strong and quite sure.

If we take into account the way music is structured, tension and relaxation represent an important phenomenon which the composer uses to create specific emotional effects on the listener. In particular, they are part of a musical process (progression) through which listeners experience a sense of expectation. Tonality itself is based on expectation; it employs a system of voice leading rules which favour a sense of tension followed by a sense of relaxation as a melody or succession of harmonies arrive at a stable tone or chord, respectively. In cases where listeners already have knowledge of Western tonal structures, they are able to anticipate the particular rhythmic, melodic or harmonic structures of the music they are listening to, mentally representing them before their occurrence.

The idea that through learning listeners become able to anticipate future events is well expressed by Titchener and Broyles (1973). In their critique of Meyer’s information theory, they make an interesting distinction between expectation and anticipation through the following quotation of Donald Sherburne:

If the [Meyer’s] theory were correct, the first hearing of a work should reek with meaning and send emotional tingles to the tips of the toes; but with subsequent hearings the significance and emotional impact ought to decline rapidly as the unexpected becomes the expected, as expectation becomes replaced by recollection and anticipation. In fact the far more common experience is that the works
Titchener and Broyles note that “Sherburne uses “anticipation” to suggest the awaiting of a known outcome and seems to contrast it with “expectation”, which suggests that the range of probable events is known but the unique event is not. Thus, when the Ideal Auditor passes from expecting to anticipating, no more information is conveyed” (1973, p. 18).

Repetition and familiarity thus allow a listener to move from the expectation of the unknown to the anticipation of something that is already known because of previous experiences. Titchener and Broyles write: “the difference between knowing and experiencing, anticipating the known and expecting the unknown, and information and information is such that each of the first terms of the distinctions requires factual knowledge on the part of the Ideal Auditor, while none of the second terms does” (1973, p. 22). In this way, the sense of expectation about possible future events is replaced by a kind of mental representation of a precise future event.

Focusing on the listener’s subjective experience, the musical structure she perceives when experiencing a musical event is determined by her learned knowledge, past experiences and beliefs. These factors also determine the creation of expectations and anticipations, which influence the constitution of her experience in different ways. The musical structure she perceives does not necessarily correspond to the notated/established one: it is a subjective matter and, for this reason, it is not possible to correctly predict the listener’s response to music, nor is it possible to give a causal account of it (as various theorists and cognitive psychologists have tried to do in relation to Western tonal music). Listeners create different expectations and anticipations in relation to the way they organize sounds, and this determines different musical experiences and musical understandings.

The difference between expectation and anticipation involves the temporal development of musical listening and has important consequences for the way people constitute musical experiences and understand unfamiliar music. I suggest that anticipation influences the probability that one will understand particular music, and I argue that one of
the reasons a listener might have a sense of misunderstanding in her encounter with unfamiliar music is that she lacks the ability to anticipate.

Considering the problem of the unexpected in music, Huron claims that “without accurate expectations, there will be no successful prediction effect. As a result, tonality oriented listening will evoke the *qualia* of confusion and discomfort” (2006, p. 344). Without anticipation, which is a condition resulting from a lack of accurate expectations, there will be a kind of failure in the subjective musical experience that causes a sense of confusion and – as we will see – misunderstanding. This is particularly common in those cases where listeners experience unfamiliar music.

Jazz Improvisation is one source of study that might be interesting in addressing the effects expectation and anticipation may have on a subject’s musical experience. When a musician performs with other musicians, she expects the music to progress in a particular way and tries to anticipate the other players’ intentions in order to allow a constant musical dialogue. This dialogue is possible because anticipation of the other musicians’ intentions allows her to have a prompt reply. Her predictions are not perfect, and because of the improvisational character of this musical practice, she must be ready to create new representations – consciously or unconsciously – about what will happen in order to reply appropriately. In such a situation, both expectation and anticipation are important for the comprehension of what is going on and the realization of a dialogue between musicians.

I will leave this problem for the moment in order to introduce another one: what happens when unexpected events occur? Form a biological point of view, the inability to create correct expectations represents a failure. When Huron underlines the biological relevance of expectation, it seems that he refers more to anticipation than to expectation, because in order for an individual or a group of individuals to survive it is important to be able to predict specific future events and react to them in the fastest possible way. Anticipation may be seen as the act of replying to an expected event before it happens. In this sense, it reflects a mechanism of response to the environment which is fundamental to survival. In the case where an event is unexpected, the subject fails because she was not able to anticipate it.
The various theories that I outlined in the previous chapter discuss expectation in relation to the perception of familiar tonal music. Based on these, we can claim that listeners learn the regularities of musical structures common to the idiom they are exposed to, and mentally anticipate a particular musical event on the basis of such learned knowledge. But the question remains: what happens with completely unfamiliar music? For example, in the case of some contemporary and atonal music, which is not very well known by the majority of people, a listener is in a state of expectation and she has a sense that something will happen during the development of the piece. She tries to create anticipations and predict future events, and she mentally represents the ongoing of the piece on the basis of her learned knowledge. But the anticipations built, for example, on the basis of tonal patterns and schemas will constantly result in wrong predictions. It follows that her musical experience will be full of misunderstanding and surprise, which is one example of an experience which falls along a continuum of different responses to the music in question.

To better understand this point, I propose a situational example coming from a non-musical context. An article about an argument that I do not know provides me with more information than it does for a reader who knows the subject of the article well. I expect that the argument will have a conclusion, but I cannot anticipate the content of that conclusion. An expert reader can anticipate the content of the article or, at least, some of the arguments it might contain. She will either be satisfied or not with the conclusion in relation to her beliefs and intentions at the moment she started the reading and she will have a better comprehension in the end. On the other side, I may be surprised by the new arguments I learn, or I may have a sense of misunderstanding because I did not have the knowledge allowing for anticipation of the article’s content. All is completely new to me, and in order

85 By ‘contemporary unfamiliar music’ I refer to those contemporary musical genres in which tonality has been replaced by other organizational systems (serial music, electroacoustic music, concrete music, experimental music, atonal music, minimalist music) and the sound experimentation has a great relevance (Contemporary classical music, post-1975, including post-modern music, Spectral music, post minimalism, sound art, etc.). Due to their lack of diffusion with respect to musical genres based on tonality, they are relatively unknown to the majority of non-musician listeners who have been brought up with Western tonal music. For this reason, these genres are perceived as non-familiars, and thus they are more difficult to appreciate and understand by untrained listeners. The same holds for music belonging to cultures that are very different from the listener’s culture. In these cases, listeners find some difficulties in the organization of sounds into musical structures. They are not able to create correct anticipations and this leads them to a sense of misunderstanding of the atonal or culturally unique music.
to better understand the content, I will likely need to read it another time. During the second reading, I start to anticipate the content, to relate it to what I have already read and, in this way, I find relationships between various parts of the article that allow me to better understand it. Anticipation is thus very important to better comprehension.

Returning to the case of the man in the metro, he organizes the sounds he hears on the basis of past experiences and knowledge. Due to specific perceptual mechanisms, he is able to create structural relationships between sounds that acquire a musical meaning. At the same time, he experiences expectations that might be followed by anticipations. This last aspect is what makes a difference to his experience: a lack of anticipations indicates that the music he hears is not familiar or it does not fit with previous concepts or beliefs. This provides some definition of the kind of musical understanding he has and also influences the meaning (musical or extra-musical) his whole auditory experience acquires.

For example, during Christmas time the man in the metro might expect to listen to Christmas traditional songs because his anticipations are based on a cultural context he is familiar with and his experience with the same cultural context in previous years. When he gets off of the metro and identifies a musical event, he may initially believe that the song he hears is a Christmas song. He will be surprised when he realizes that it is actually an Irish traditional song. In a case where the music is completely unfamiliar, he will create expectations that are not satisfied and thus he will be incapable of anticipating musical events. For this reason, he has a sense of misunderstanding and confusion.

To conclude, in this chapter I have tried to focus on some arguments that support the conceptual distinction between expectation and anticipation. I assert expectation as a mental state of suspense about what is going on, during which a range of possible events are expected to happen. Expectation contributes to the constitution of meaningful subjective experiences and the realization of the intentional movement that allows a listener to be conscious of music and of a sequence of sounds as having musical meanings. Anticipation may be seen as a particular moment in the process of expectation: it is the moment in which the listener mentally predicts how the music will go on. More precisely, it is the ability to have a kind of mental representation of a future event, which influences the listener’s
comprehension of music and ultimately may define the difference in her understanding or misunderstanding of unfamiliar music.

In the next chapter I will consider an assumption that a lot of empirical studies make about perceived musical structure: the idea that the perceived structure corresponds to the one that has been notated by the composer in the score. The aim is to underlie the subjective character of musical experience. I will then develop the distinction between expectation and anticipation through the analysis of the temporal structure of human consciousness. In particular, I will try to understand how they influence basic musical understanding and, therefore, the constitution of subjective musical experiences. I will also question the way anticipation is to be understood: is it a kind of perceptual experience, a mental representation or an act of imagination? Is it developed at a conscious or unconscious level?
THE PERCEIVED MUSICAL STRUCTURE VERSUS
THE NOTATED/ESTABLISHED MUSICAL STRUCTURE

Introduction

In the first and second chapters I introduced the problem of how listeners understand an acoustic event occurring in the auditory environment as musically meaningful. I also introduced the concept of ‘basic musical understanding’ to describe the immediate perceptual awareness of music, and related it to the process of constituting subjective musical experiences. A fundamental aspect of such a process is the organization of sounds into structural forms, which influences the subjective meaning an auditory experience acquires. Thus, the question is: how can we describe the structure that listeners perceive while hearing music?

Some studies in the cognitive psychology of music try to explain the processes through which musical structures are perceived, but they often presuppose that the perceived structure corresponds to the one that has been notated by the composer in the score. This is part of a common view held by those familiar with the Western Classical tonal tradition. From this perspective, a piece of music finds its identity in the score and listeners understand music when they grasp its structural features. Studies based on such an assumptions may shed some light on certain perceptual processes, but do not describe the real perceptual experience of listeners because they do not take into account the different factors, contexts, situations and types of music that influence one’s way of hearing and understanding it.

On the basis of the issues previously presented, in this chapter I try to distinguish between perceived and notated/established structure, showing that in order to constitute musical experiences at the perceptual level it is not necessary to organize sounds according
to the structure imposed by the composer or the performer. I support this idea by taking into account the role of expectation in the structural organization of sounds.

**Listeners perceive music structurally.**

*The notion of ‘structural hearing’.*

We have widely seen that the organizational aspect of perceptual activity is fundamental in order to understand the sound environment as music and to constitute musical experiences. The perceiver has the natural tendency to look for relationships between sounds and to find meaningful forms while hearing sound events occurring in the acoustical environment. Most empirical studies – as well as studies in music theory – support the idea that such a tendency is realized through structural and hierarchical organization. In other words, the human mind works structurally in order to understand the auditory environment and grasp forms, which may be musical or otherwise. The man in the metro, for example, structurally organizes the acoustic environment in order to understand what is happening around him. This allows him to perceive music.

Structural organization is important to the organization of sounds into a unified, coherent whole. Listeners unconsciously look for wholeness, and the relationship between part and whole – which is fundamental to the constitution of meaningful musical experience – is realized through structural and hierarchical organization. At the basis of such organization there are certain processes going on which are unique to each individual, but the tendency to look for structural relationships is shared by all perceivers. For this reason it may be considered a natural tendency of the human mind. As Levinson has suggested, listeners understand music while following its temporal development and relating sounds into a coherent structural unit. The structural relationships that they find between sounds are fundamental to the lowest level of musical listening, in the moment-to-

---

86 For a brief summary see Levinson (2006).
moment movement of the musical flux. This is not carried through at the high level of the musical organization – which often involves intellectual musical understanding – but is realized at the perceptual level.

As many empirical studies make clear, listeners organize music according to the musical patterns and schema that they have consciously or unconsciously learned during their musical experiences or through simple exposure to the music of their own culture. Such experience and knowledge determines different ways of organizing music, particularly between musicians and non-musicians.

A lot of studies have tried to explain sensitivity to structure in music. For example, in the work *Exploring the Sensitivity to Structure in Music* Annabel Cohen (1982) describes two contrasting approaches to the acquisition of sensitivity to musical structures that were applied in experimental psychology of the empiricist tradition and in Gestalt psychology during the past century. One such approach involves learning and the other involving common human abilities. In her paper, Cohen presents research which illustrates the interaction between learned and innate human abilities and the potential for integration. From this perspective, the listener’s sensitivity to a musical structure is in part dependant on exposure to a particular musical culture and the acquisition of its typical recurrent structures, and in the other part reflects a natural tendency of the human mind (as the Gestalt principles describe).

As some studies have demonstrated, such a tendency to organize sounds into structural forms (or following the structural features of music) is evident even in infants’ behaviour and responses to music. One such interesting work is Krumhansl and Jusczyk’s: *Infants’ Perception of Phrase Structure in Music* (1990). Through a visual procedure and the head-turn preference procedure, they examined 6 and 4½-month-old infants’

---

87 As Huron observes, “the statistical properties of the sequence are learned as a by-product of simple exposure without any conscious awareness by the listener” (2006, p. 71).
88 For an overview see Schellenberg & Trehub (1999); Krumhansl & Jusczyk (1990); Trehub (1987); Schellenberg & Trehub (1996); Trehub & Thorpe (1989); Demany, McKenzie & Vurpillot (1977). These studies show the human tendency to organize music in a structural way. Changes in the way music is structured influence music perception and this indicates a particular sensitivity to the structural aspect of music. Structural organization is thus fundamental to the musical experience.
89 The head turn preference procedure is a conjugate reinforcement technique used to collect behavioural data from infant subjects. It involves teaching the infant that when they turn their head in a certain way, usually to
sensitivity to phrase structure in music and tried to identify the factors underlying this sensitivity. They used the initial sections of Mozart minuets (because of their regular melodic, rhythmic and harmonic structure), which were divided into segments that either did or did not correspond to the phrase structure of the notated music. As the figure below shows, some versions (called Natural) were divided by short pauses into segments that corresponded to the original phrases in the music (Krumhansl & Jusczyk, 1990, p. 71). In other versions (called Unnatural) the pauses were inserted in the middle of the phrase, thus changing the original structure of the minuet.

![Figure 5. Natural and Unnatural versions of the initial section of a Mozart’s minuet.](image)

The results showed that the infants preferred the Natural segments, indicating their ability to segment and group musical passages and, therefore to organize sounds into face a visual stimuli, an auditory stimulus will become activated. The head-turn preference procedure is used to analyse sensitivity to various aspects of music through the investigation of a child’s preference for one version of the musical excerpts presented during an experiment. Such a preference is established by measuring the time spent listening to different musical stimuli. More precisely, this was accomplished by beginning the Unnatural version at some point other than the beginning of the piece while keeping the pattern of inserted pauses unchanged.
structural forms. Their behaviour also suggested that protracted musical experience may not be necessary to perceive phrase structure in music.

It follows that the perception of sounds as structurally organized is fundamental for the constitution of musical experiences. As Isabelle Peretz observes, changing the structural features of music influences the way listeners perceive music and the meaning it acquires for the subject. Referring to the emotional meaning that a musical experience can have, she claims that “in music it is easier to manipulate structural characteristics systematically so that the same music can express different emotions gradually” (Peretz, 2001, p. 122). However, the fact that a structural change can influence the listener’s emotions does not necessarily correspond with a change in the perceived structure.

As noted by Blacking (1986), music is ‘humanly organized sound’ and, in Scruton’s words, “whatever it is, music is not a natural kind” (1997, p. 16-17). The existence of music is dependant on the way human beings organize sounds and interact with the auditory environment. In the case of music listening, sounds are perceptually patterned by listeners in order to develop musical structures. As Daniel Levitin, for example, claims, the tonal key “has to do with a hierarchy of importance that exists between tones in a musical piece; this hierarchy does not exist in-the-world, it exists only in our minds, as a function of our experiences with a musical style and musical idioms, and mental schemas that all of us develop for understanding music” (2007, p. 17).

The notion of ‘structural hearing’ refers to the idea that the human mind works structurally in order to organize sound events occurring in the acoustical environment into musical forms. It describes the basic and fundamental way of hearing and organizing music which, in my opinion, guarantees the subjective aspect of the perceptual musical experience and allows the development of different ways of listening.
The concept of ‘structural listener’.

A concept which is related to the notion of ‘structural hearing’ is that of the ‘structural listener’. Traditionally, this refers to the problem of how an ‘expert listener’ may be described and defined. In Theodor Adorno’s *Introduction to the Sociology of Music* (1976), for example, ‘expert’ and ‘good’ listeners are defined in relation to the specific quality of what is heard. A first presupposition of his view is that musical works are objectively structured and have meaning in themselves. Such structures can be perceived and experienced with different degrees of accuracy. The *expert listener* is “the fully conscious listener who tends to miss nothing and at the same time, at each moment, accounts to himself for what he has heard. […] Spontaneously following the course of music, even complicated music, he hears the sequence, hears past, present, and future moments together so that they crystallize into a meaningful context” (1976, p. 4). Such a listener develops a type of structural hearing which follows the musical logic and the technical elements of the composition. The way she listens to music indicates an intellectual understanding based on her specific musical education and training.

From Adorno’s perspective the expert and good listener are distinct: “the good listener too hears beyond musical details, makes connections spontaneously, and judges for good reasons, not just by categories of prestige and by an arbitrary taste; but he is not, or not fully, aware of the technical and structural implications. Having unconsciously mastered its immanent logic, he understands music about the way we understand our own language even though virtually or wholly ignorant of its grammar and syntax” (1976, p. 5). Moreover, “this is the type we mean when we speak of “a musical person” – assuming that this phrase is still a reminder of the capacity for direct, meaningful hearing, that it is not enough for us to find that someone “likes” music. Historically, such musicality required a certain homogeneity of musical culture” (1976, p. 5-6).

---

91 Basically, the concept of ‘expert listener’ refers to the listener who has acquired particular musical knowledge and organizes music on that basis. Meyer, Levinson, Kivy, Lerdahl and Jackendoff among others use this concept.
A lot of theories and empirical studies (such as those presented by Meyer, 1956; Bigand, 1990a; Lerdahl & Jackendoff, 1983a) make reference to this kind of listener. This is clear if we consider the relevance that learning and exposure have for musical listening. Adorno’s definitions of ‘good’ and ‘expert’ listeners follow the idea that, in order to understand a piece of music it is necessary to follow an acquired musical syntax or grammar (as with language), even in the case where a listener is not aware of a piece’s musical structure while she is listening to it. But music and its meaning is not as constrained by grammatical rules as language is, and listeners might follow different logics and categories in their musical hearing.

I suggest that the concept of a ‘structural listener’ simply sums up the idea that the perceiver’s mind works structurally in order to find relationships between sounds and to constitute musical experiences. The use of a musical syntax and the ability to understand the technical aspects of a musical composition indicate different degrees of knowledge and understanding.

**Assumption about the perceived structure.**

**Preliminary conceptual distinctions.**

I propose a distinction between two ways of conceiving musical structure which are related to different musical experiences. The first one corresponds to the way a composer has organized sounds and notated them in the score during the act of creating a musical work. It can also be identified with the hierarchical organization that a performer establishes (consciously or in a spontaneous way) during her performance or improvisational practice. In both cases this refers to a sound organization which is imposed by someone with the intention of creating music and making it available to other people (and, thus, sharing a musical experience). I call this kind of structure the ‘notated/established structure’.
If we consider the listener’s perception of music, the structure of a musical piece corresponds to the way she organizes sounds. In the present work, I focus only on the second way of conceiving musical structure, which concerns the perceptual experience and the way listeners unconsciously organize sounds in order to be aware of their musical experiences. Such an organization is fundamentally a structural one, and it reflects the way the human mind works and acts in the living world. This refers to a way of hearing music in which listeners are not merely passively affected, but rather they are actively constructing their own ability to be moved. I call the result of this process the ‘perceived structure’. In this case, perception is a creative and generative process.

The perceived structure differs from the notated/established one not only because it refers to a very different musical experience (perceptual listening versus compositional and performing experience), but it results in a subjective musical experience which opens up the possibility for different interpretations and musical understanding.

A problem with a lot of theories and studies about music perception is the presupposition that the perceived structure corresponds to the notated/established one. This is related to the idea of ‘correct’ musical understanding, for which the perception of the notated structure is seen as a condition. Ideas based on this presupposition, however, fail to thoroughly explain the way human beings perceive music – particularly with respect to unfamiliar music.

Comparing the notated structure of a piece of music to the way listeners perceive it in real contexts (like the case of the man in the metro) is a difficult task, but one which may be important in order to better comprehend music perception. The way in which melodies are notationally structured does not necessarily coincide with the way listeners (even experienced listeners) perceive them. Thus, a comparison of this sort will allow us to recognize the differences between perceived and notated structure, which will shed some light on common perceptual additions, subtractions and manipulations.

Another clarification needs to be made if we want to understand why the perceived structure does not necessarily correspond to the notated/established one. It concerns the distinction between ‘perceptual understanding’ and ‘intellectual understanding’ of a piece’s
musical structure. The comprehension of structural models and the awareness of the correct structure (the notated/established one) are part of an intellectual understanding which requires reflection and a conscious process of abstraction of the perceived structure. Intellectual understanding may be defined as the ability to comprehend high-level structural relationships and large-scale forms; the awareness of the architectonic structure of a musical composition which, in Kivy and Hanslick’s opinion, allows a full appreciation of it.

According to Cook and Levinson, the perceptual act of listening to music has less to do with large-scale structural models and more to do with the multitude of local, intuitive decisions that give rise to the music that is heard. What is most important in defining the structural aspect of a piece’s organization is the ability to create relationships between sounds in a way that results in a number of small musical parts being combined into a unitary whole. As Koopman and Davies claim, “relationships between parts of a musical work are established simply by their being concretely perceived as belonging together within the same perceptual fields” (2001, p. 263).

Another important idea is that understanding what is going on at the perceptual level implies the ability to relate sounds in the moment. The basic perceptual understanding of a musical event might be attained without any awareness whatsoever of the overall structure of a musical piece. In this sense, it is developed in the moment-to-moment unfolding of a piece of music. I will develop this idea in the next chapter.

The experience of structure at the perceptual level is fundamental in order to organize sounds and understand them as constituting music, but it is a kind of musical comprehension which is very different from intellectual understanding of the structural elements of a musical composition. The case of the man in the metro refers to a perceptual understanding, which is the focus of my work.

_Scruton and the ‘tonal structural hearing’._

A theory which highlights the relevance of structural hearing has been developed by Roger Scruton in his book _The Aesthetic of Music_ (1997). In his opinion, listeners hear
music and its features (such as harmony) because of a structural organization of tones. He presents the idea of structural hearing as fundamental in order to articulate parts into a whole and understand music within the context of a Western tonal framework. He claims that “understanding music is in part a cognitive activity: an activity of mental organization, which collects sounds together and registers them as tones, arranged in a tonal order” (1997, p. 211).

From this perspective, understanding music is a fact of experience and is related to a specific way of hearing. Scruton relates such a way of hearing music to the tonal system, using tonality as the basis for a persuasive theory of what listeners hear. The idea is that listeners understand a piece of music because they grasp the tonal order which generates its musical surface. In this way, he gives a great deal of relevance to tonality, which is considered not just a style, but an order that listeners hear in music despite the vast divergences of style. It follows that when we describe the tonal order of a piece we are describing what is heard, when it is heard as music. Scruton supports the idea that listeners ‘naturally’ hear music in a tonal way. In his opinion, the Gestalt theories, as they relate to the organization of music, tend toward the direction typified by tonal music. This, according to Scruton, is the reason listeners cannot understand atonal music: they listen to it following a tonal structure and syntax. Atonal music is thus criticized because it does not reflect the way individuals naturally listen to music. He tries to support this idea with the example of the cadence, which reflects a sense of closure, of an end, which is an integral part of the experience of structure and form. Music lacking this structural element, such as atonal music, thus remains unclosed and listeners are not able to understand it.

Within the framework of Scruton’s aesthetic theory of music, the main idea is that listeners hear sounds as tones belonging to a metaphorical tonal space. This space is characterized by specific structural relationships that are determined by a musical syntax.

93 Scruton supports this idea by observing: “if you imagine what music in our tradition would be without the cadence, and without the ‘sense of an ending’ that comes with it, you will surely agree that the cadence is not an embellishment, but an integral part of the experience of structure and form. Atonal music can create the sense of an ending – but it is often not by the inner logic of the music, which remains ‘uncolosed’ even beyond the final measure” (1997, p. 256).
Making an analogy with language, he supports the view that music is a rule-governed art, whose meaning is worked out through its structure. Tonality is seen as a language through which listeners understand music: it is characterized by a syntax, or a system of grammatical rules that determine the structural organization of sounds and the meaning that the resulting structures acquire. It follows that the meaning of music lies in its structure, as this quotation makes clear: “it was supposed that the artistic meaning of a work is to be found ‘encoded’ in its ‘structure’, to be recuperated (perhaps at an unconscious level) by the reader who grasps the structure and identifies its elements. The key assumption was that meaning and structure are connected” (Scruton, 1997, p. 173).

Another important aspect of structural organization is that the structural features underlying the identity of a piece of music can be found in the score. It is a design that reflects the composer’s intentions, which have to be respected during performance in order to maintain the identity of the work. This leads to the idea that understanding music corresponds to the ability to grasp its notated structural pattern. In the case of listening to the performance of a musical work, Scruton argues that “to identify the work of music in the material world is to identify the sound pattern intended by the composer, which is realized in performance by producing sound events. This sound pattern defines the salient features of the musical work, and can be written in the form of a score” (1997, p. 109). According to this point of view, it follows that a work of music can be fully identified through a notational system.

All of this is related to the assumption about ‘correct’ and ‘incorrect’ musical understanding. Listeners understand music when they perceive it correctly, and to accomplish this task they must hear music in accordance with its notated structure. Moreover, they have to understand the meaning that a piece of music has within the

---

94 This view is supported by Scruton, who writes: “the works of music, whatever they are, originate in human actions, and are understood as intended objects. The design which determines the performances of a work of music is an intended design, and the intention is underdetermined by the score which records it. Whether we could an arrangement as a version of the original or as a new work, will depend in part on the intention of the arranger. And the difference between a performance and a travesty lies in our sense of the distance between the composer’s intention and the performer’s product” (1997, p. 107). The composer’s intentions thus influence listeners’ perception.
musical practice and aesthetic theory of which it is part. A meaning which, according to Scruton, depends on a metaphorical hearing that is based on tonal structure.

The following quote by Davies presents an example of structural listening that well explains the idea of ‘correct’ understanding:

she will consider what is expressed in the music in relation to more obviously structural elements. In general, she has to follow the progress of the music, so that its course makes sense. This requires of the listener that she be able to predict how the music will continue at any given moment, that she is often right in her predictions, and that where she is not, she is able to distinguish between an unanticipated but appropriate continuation and a performance or compositional blunder. Typically, she can identify performance errors as such, and she often knows what should have been played instead. She experiences the music as unfolding in a “logical” way, with what preceded as justifying or making appropriate what follows (2007, p. 6).

This is based on the presupposition that, in order to understand music, the structure perceived by the listener has to be the same as the notated one.

This is the main presupposition that I question, with particular reference to the problem concerning the constitution of subjective musical experiences. In my opinion, such a presupposition does not allow for a comprehensive description of the real perceptual experience of music, particularly in those cases in which listeners are not very familiar with the music that they hear. Moreover, it does not take into account the distinction between different kinds and levels of musical understanding: as we have already seen, the perceptual understanding is different from the intellectual understanding of a musical piece. Finally, it contains a specific ontological point of view, for which the identity of a musical work has to be found in its notated structure.95 I do not take this last problem into account, as to do so would require a new discussion concerning musical identity. In outlining these ideas, I only wish to point out that Scruton’s perspective is too narrow in that it is inextricably tied to the tonal classical tradition and does not acknowledge different ways of making and hearing music. In his view, tonality seems to be a human universal: the way all human beings hear and understand music.

95 The problem of musical identity has been widely discussed within the philosophical debate concerning the ontological status of works of art. See, for example, Davies (2001b; 2003); Dodd (2007); Goehr (1992); Goodman (1968); Levinson (1990); Predelli (2001); Rohrbaugh (2003); Wolterstorff (1980).
In my opinion, views held by scholars such as Scruton do not explain the constitution of different subjective experiences. Firstly, a great deal of music exists which is organized differently than Western tonal music. Organization based on tonal music theory and notation is only one way of hearing and structurally organizing sounds, and it may not correspond to the perceived organization of non-musician listeners. For this reason I prefer the idea that the structural order listeners hear between sounds is the result of a learning process which depends on the listeners’ experiences. Music, in this sense, depends upon a living musical culture.

Some other relevant theories.

In their *Generative Theory of Tonal Music*, Lerdahl and Jackendoff (1983a) support the idea that listeners understand a piece of music when they grasp its structural features. Their theory is also based on Western tonal music and gives particular attention to Schenkerian structural analysis and Chomsky’s generative linguistic theory. In looking for the cognitive principles that explain the comprehension that expert listeners have of tonal music, they refer to Gestalt psychology but do not present empirical studies to support their arguments. Thus, they build their generative theory only on the basis of theoretical assumptions and do not take subjective perceptual experience or the real contexts in which listeners have musical experiences into account. In particular, the two authors underline the hierarchical aspect of tonal music as fundamental for the organization of sounds into structures and, therefore, for the comprehension and identification of music.96 Moreover, they describe the structure that the experienced listener infers from the musical surface as reflecting the way a musical work is organized in a score. According to these ideas, a listener understands the work only when she cognitively grasps the notated structure. This ability allows her to recognize the quality of a performance of a piece of music.

---

96 Lerdahl and Jackendoff describe hierarchical musical structure in the following way: “a hierarchical structure, in the sense used in this theory, is an organization composed of discrete elements or regions related in such a way that one element or region subsumes or contains other elements or regions. […] In principle this process of subordination (or domination) can continue indefinitely” (1983a, p. 13).
Meyer has also suggested that understanding music is related to the comprehension of its structure.\textsuperscript{97} From his perspective, both the perceptual organization and the creation of expectations are related to tonal structure. His theory implies the idea that the creation of expectations follows the structure of the music one is listening to. It is based on tonal classical music, which is highly hierarchically structured and grounded on the creation of expectations. Hanslick (1957), who argues that music is beautiful because of its structural composition, follows the same idea when he claims that the value of music is in its structure and that it can be fully appreciated only if such a structure is understood. It is clear that these theories all assume the listener’s ability to detect the notated/established structure during the act of perceiving a piece of music and that it this ability is fundamental to a ‘correct’ musical understanding.

\textit{Empirical examples of structural hearing.}

Some of the research conducted by Bigand and his colleagues is very useful to this discussion in that it allows us to analyse the idea that listeners understand and identify a piece of music when they grasp its structure. As we will see, the principles of recognition and familiarity are relevant concepts within this perspective and they follow the idea that listeners understand music when they are familiar with its structural features. In the article \textit{Abstraction of Two Forms of Underlying Structure in a Tonal Melody} (1990a), Bigand describes an experiment based on Lerdahl and Jackendoff’s theory which was designed to test the abstraction of two forms of underlying structures in a melodic passage: ‘reduced structure’ and ‘prolongational structure’. During the experiment he firstly presented participants with a family of four melodies, all of which had the same underlying structure but different rhythmic-melodic organization. Successively, he interspersed these melodies with four analogous melodies differing in their underlying structure. participants were asked to identify melodies belonging to the first family. In order to investigate participants’ abilities, Bigand presented either a real family or a false family. In this way, he was able to

\textsuperscript{97} As we have already seen, Meyer explains musical meaning in terms of relationships between musical elements. From such a perspective, explaining meaning means translating it into music-structural terms.
show that the subjects (both musicians and non-musicians) succeed in abstracting underlying structures.

**The Experiment**

Figure 6. The two families of melodies (Bigand, 1990a, p. 49).

One of the aims of his experiment was “to work out how one gets from the perception of the musical structure, made up of sounds of differing pitches, durations, intensities, timbre qualities, to the extraction of underlying harmonic structures, indispensable for proper understanding of tonal pieces” (1990a, p. 47).
The internalisation of hierarchical relationships between notes is culturally influenced and highly dependent on the process of reduction. In the case of tonal music, for example, listeners are able to detect hierarchical structures through reductions which separate ornaments from what is structurally important. Another important factor that influenced the listener’s behaviour is their individual memorisation of melodies. The results of Bigand’s experiment show that a listener is able to abstract an underlying structure common to four melodies that have different rhythmic-melodic contours. Moreover, she can distinguish between two families of melodies with comparable surfaces, on the basis of their underlying structures. These results demonstrate that a listener can go beyond the level of superficial organization to reduce the musical surfaces to a minimal underlying structure representing them. The listener, then, manages “to hear melodies belonging to the same family as variations on an underlying pattern which is more important than the differences observed on the musical surface” (Bigand, 1990a, p. 57). This aptitude to abstract underlying structures is crucial for the understanding of tonal music “since, without it, it would be impossible, for example, to grasp the relationship between a theme and its variations, nor to appreciate, on the other hand, the specific contribution of each individual variation” (1990a, p. 58). It would also be impossible to recognize different melodies as belonging to the same family.

In this case, the identification of music is related to a process of abstraction that makes the listener familiar with particular musical structures and allows her to recognize them if they occur again. Familiarity refers both to the past experience of a musical phenomenon that she retains in memory and to the acquisition of recurrent structures in the music she is most exposed to. Recognition is related to the ability of the listener to relate a new musical event to something she is familiar with and to discriminate its features in order to judge it as the same or different from the previous one. In the example of Bigand’s experiment, subjects recognize melodies as the same (as belonging to the same family) because they abstract and become familiar with their underlying structures. An interesting

98 In Western tonal music theory, ornaments are musical flourishes that are not necessary to carry the overall line of the melody (or harmony), but serve instead to embellish or ‘ornament’ the line. Many ornaments are performed as ‘fast notes’ around a central note. The amount of ornamentation in a piece of music can vary from quite extensive (it was often so in the Baroque period) to relatively little or even none.
aspect of this study is that the majority of the participants, regardless of musical experience, succeeded in the tasks presented during the experiment. The fact that both musicians and non-musicians were able to recognize melodies with the same underlying structure as belonging to the same family suggests that their perceptual activity works in the same basic way.

Bigand identifies three factors that determine music perception and allow the passage from the perception of a musical surface to the comprehension of its structures:
- knowledge of the hierarchy of a musical system (knowledge of the rules used in a musical system, that could be conscious or unconsciously acquired, general or very specific);
- principles of perceptual organizations;
- cognitive constraints involved in the development of more complex relationships.

These factors allow a listener to find the fundamental structure of a piece of music and, according to the idea that the meaning of a piece of music resides in its structure, to grasp the meaning a composer intended to communicate.

The first problem with Bigand’s approach concerns the definition of the elements that determine the recognition of a piece of music and that characterize its identity. The second issue is that it presupposes that listeners (both musicians and non-musicians) perceive the same tonal structure. The fact that they are able to recognize melodies as belonging to the same family does not necessarily imply that they organize music in the same way or according to the notated structure. Moreover, how can we be sure that such recognition is realized thanks to a structural organization rather than the recognition of other relevant features? Finally, does the fact that the participants in Bigand’s experiment were able to recognize families of melodies allow us to say that they naturally hear music in a tonal way – as Scruton argued – or do they learn how to recognize melodies?

In making reference to tonal classical music, which is highly structurally organized, studies like those of Bigand (as well as Lerdahl & Jackendoff, 1983b; Deutsch, 1982a, 1982b; Deliege, 1987) relate music comprehension to the fundamental and underlying structure of a piece or sequence of music. Listeners understand music when they find,
through a system of rules that constitutes their musical knowledge and through specific
cognitive processes, the hierarchical relationships that characterize the structure of a piece
of music. In this sense, comprehension is the ability to process the ‘sous-jacent’ – that is,
the ‘underlying’ – structure of music. This is in line with a formalist point of view and the
aesthetic theories of music presented by philosophers like Kivy and Scruton.99

Without taking into account the problem of how the underlying structure has to be
understood, an interesting point to highlight for the purposes of the present discussion is the
assumption that listeners, in hearing music, follow the tonal structure that has been imposed
to a melody or family of melodies. Under this assumption, tonal structure and its
comprehension by the listener is the condition for a ‘correct’ musical understanding. In my
opinion, the problem of this perspective is that it does not take into account the different
ways of hearing music that can be developed by listeners, and it confines perceptual
understanding to a ‘correct’ structural organization. This neglects a kind of basic perceptual
understanding for which it is not necessary to grasp the notated structure in order to
constitute musical experiences. The man in the metro, for example, immediately perceives
music even if he does not grasp its notated structure (if any). Moreover, this approach does
not allow the listener to have music comprehension in cases where she listens to unfamiliar
music.

The case of transposed melodies.

Some empirical studies about the recognition of transposed melodies or the
detection of structural changes seem to support the assumption that the perceived structure
corresponds to the notated/established one. The underlying structure to which they give
particular attention corresponds to a compositional model which is notated in the score and
embellished by ornaments and variations. For example, studies about infants’ musical
behaviour made by Glenn Schellenberg and Sandra Trehub (1999) show that the ability to

99 See Scruton (1997); Kivy (2002). Levinson (1997) also follows the idea that music comprehension (even at
its basic perceptual level) is related to the ability to grasp its structure.
grasp musical structures may be largely innate, since the infants were responsive to musical structures characterized by particular stability.

The figure below presents two series of musical sequences that include the transposed repetition of the first standard sequence and sequences with subtle variations to the first series and significant structural changes to the second one (Schellenberg & Trehub, 1999, p. 114). The fact that infants turn their heads when the structural change occurs indicates that they are able to detect modifications to the underlying fundamental structure. It follows that they are able to recognize musical sequences which they have previously identified. As in the case of Bigand’s experiment, the identification process is related to the processing and individuation of underlying structures.

![Musical sequences](image)

**Figure 7.** Repetitions of a standard sequence with or without changes.

In her work *Organizational Process in Music* (1982a), Diana Deutsch first analyses a series of experiments that apply Gestalt principles to music and then she considers the high-order musical abstraction which guides equivalences and perceptual similarities. In particular, she tries to explain how a listener is able to recognize the form of a musical sequence when it is altered in a way that does not modify its underlying structure and the fundamental relationships that characterize its internal organization (because these seem to be the elements that guarantee its identity and enable recognition). The Gestalt psychologists emphasized the fact that the same musical configurations may be perceived as equivalents when they are entirely composed by different elements but the fundamental
relationships between such elements are preserved. A visual form, for example, maintains its perceptual identity even when it is reported from different places in the visual field or when it is altered in size and presented as it’s a mirror image because of the fact that it preserves its fundamental structural relationships.

One of the first examples of this perceptual equivalence (that shows the ability to recognize previously identified structures and forms) was made with reference to the perception of music. Christian Von Ehrenfels (1890) pointed out that a melody, when transposed, preserves its essential forms and the fundamental relationships between the tones by which it is composed. The behaviour of the melodies is thus the same as that of the visual forms presented in the previous example, which do not change when they are put in another place or position: they are still recognizable.

The following figure clearly shows this behaviour (Deutsch, 1982a, p. 127).

![Figure 8. Equivalences in auditory and visual perception.](image-url)
This recognition activity shows the capacity to detect the underlying structure of a melody even if variations or modifications have been made at the superficial level. This may also be understood as a characteristic of the perceptual process which is fundamental for the ‘correct’ comprehension of music.

In my opinion, these types of experiments are interesting in that they show the perceiver’s attention to the structural aspects of music, supporting the idea of structural hearing. However, they do not prove an ability to detect the ‘underlying’ musical structure: this is merely a suggestion that has been made by various theorists in attempt to explain music perception and recognition on the basis of the notated structure. This approach masks some difficulties. Firstly, relying on the analysis of music perception based on notation alone does not take into account the fact that music from a theoretical point of view differs from music considered from a perceptual point of view. Thus, music which has been transposed could be understood in the sphere of music perception quite differently from the way it is understood in the sphere of music theory. The assumption that listeners are able to perceive the notated structure of music may be wrong.

Another problem that theorists subscribing to this point of view do not take into consideration is that ‘perceiving as the same’ is different from ‘recognizing as the same’. Recognition of melodies can be made in relation to various contextual and experiential aspects. This implies that it is not necessary to perceive the same structure (as it is notated or previously perceived) in order to recognize the same piece of music.

Finally, these theorists do not question the influence of other important musical dimensions on the perception and structural organization of music. For example, what happens in the case of timbre changes? The man in the metro immediately recognizes that the song a busking guitarist is playing is the same one which Leonard Cohen played on the piano during one of his recent concerts in Montréal,— that is, *Hallelujah*. In this case, his recognition might be due to different factors and abilities. According to the formalists, he might recognize the guitarist’s performance in the metro as corresponding to a specific song because of the melodic line, but he may also recognize it on the basis of the words

---

100 In Deutsch (1982a, pp. 126-127) some experimental data are discussed in order to investigate this ability.
used or other superficial dimensions. Moreover, he may recognize the two versions as the same in relation to previous encounters with the song and on the basis of his perceptual abilities. The harmonic structure of the guitarist’s version might be different from the piano version, but the work is still recognizable and maintains specific meanings for the subject who experiences it (in this case, the man in the metro). In addition to these considerations, it might also happen that the man in the metro does not recognize the guitarist’s versions as corresponding to Cohen’s song because of the timbral change. These points suggest that music comprehension and the constitution of musical experiences are not necessarily tied to the perception of the ‘correct’ notated structure, which therefore does not coincide with the perceived one.

**Perception of musical structure: critical analyses of several proposed theories.**

In this second part of the chapter, I try to show that the theories outlined thus far – as well as the formalist’s point of view, which is supported by those who believe that listeners understand music in terms of its structural organization – do not account for the complex perceptual experience that a listener has in real world contexts. Although such studies are very important in providing new perspectives about the way the human mind works and in helping us to understand how the phenomenon of music becomes possible, they are too focused on the Western classical tonal tradition and do not take other musical genres into account. The presupposition these theories make is a limit to the constitution of meaning and it neglects the possibility that listeners can have musical understanding even if they do not grasp the ‘correct’ notated structure.

As I aver out at the beginning of this chapter, listeners perceive and organize music structurally according to specific cognitive mechanisms and learned schema. But listeners organize sounds into musical structures in different ways. This determines different ways of listening to music and the constitution of different musical experiences. I follow the idea of
music as ‘humanly organized sounds’ (Blacking, 1973), but unlike formalists and structuralists, I do not believe that in order to have basic perceptual understanding and to constitute musical experiences in a complex auditory environment listeners necessarily have to grasp the notated/established structure.

The aim of the work *The Angel of Death* (McAdams, Reynolds & Battier, 2005), for example, is to see whether listeners actually hear what composers intend them to hear, (and what musicologists think they will hear). Although experimental studies conducted by the authors did provide evidence that there are connections between a composer’s intent and a listener’s understanding (thus supporting music as a form of communication), the researchers found that the structure which the listeners perceived was in fact often different from the one the composer tried to convey through her work. One reason for this difference lies in the complexity of the musical materials used. However, despite the fact that the musical excerpts presented in the experiment were often unknown to the listeners, they were still able to organize them on the basis of superficial musical qualities and through cognitive principles such as similarity. I will return to this topic in the seventh chapter. For the moment, I suggest that the subjective aspect of the structural organization of sounds might also be revealed also by studying meter perception, which is subjectively determined. From this perspective, the causal account of the listeners’ perceptual experience (which certain authors have tried to develop as a focus of their writing\(^{101}\)) is put into question and it becomes difficult to support with certain kinds of music – in particular those that are unfamiliar to the listener and which lack of a precise notational system.

*The Structural way of hearing music as a learned process.*

The theories and works that I have presented thus far follow the hypothesis that musical listening is based on structural hearing which is essentially tonal. They also try to give a causal account of the perception of tonal music. Meyer, for example, argues that

\(^{101}\) Meyer, for example, gives a causal account of the listener’s affective experience in looking for causal relationships between the structure of a musical piece and the expectations such relationships arise in the listener. In this way he tries to explain the listener’s subjective experience by considering the music itself.
listeners create expectations based on tonality, and this allows the explanation of music perception in terms of causal relationships between sounds: following the tonal structure, a sound implies the following one and so on. In this way the compositional practice is supported and justified in its attempt to build structures which will have a specific effect on listeners.

One of the most relevant phenomena characteristic of tonal music is the tendency toward some sort of closure. Musical closure represents both a perceptual tendency – corresponding to the Gestalt principle of good continuation\(^\text{102}\) – and a syntactic concept that characterizes the movement of tension and relaxation in tonal music. A typical structural element of tonal music that indicates a resolution from tension to rest is the cadence. It exemplifies the Western tonal listeners’ sense of tonic closure and is considered an integral part of the experience of structure and form. Some authors like Scruton view this as an example of the natural way in which people hear music and organize sounds. Scruton observes that in the case of atonal music, whose inner logic does not contain the cadence, music creates in the listener some sense of an ending which remains ‘unclosed’. The fact that listeners have this sense of closure upon hearing a cadence does not, however, lead to the conclusion that they hear music tonally or according to tonal structures: this is, in my opinion, only a presupposition that does not take into account more complex phenomena. It reflects the intentional movement of human consciousness, but the way it is determined depends upon auditory learning and subjective experiences.

Through tonality, musicians of the 18th century tried to express interiority through conventional structures and syntax. Susan McClary writes: “musicians believed that they had finally perfected a vocabulary for representing the universals of subjective feelings within a universally accessible system of social encoding” (2000, p. 72). But what happens in the case of music which is not tonal? In Scruton’s opinion, in such cases listeners try to hear the music tonally because tonal structural hearing is something like a ‘universal’ way

\(^{102}\) In the field of visual perception, the Gestalt principle of ‘good continuation’ states that graphic elements which suggest a continued visual line will tend to be grouped together. In addition, visual patterns with good continuation may suggest to the viewer that the pattern continues beyond the end of what they are able to see of the presented visual stimulus. That is, we mentally ‘fill in’ or ‘paint in’ the rest of the pattern.
of listening to music, or the natural way of organizing sounds. This leads to the conclusion that listeners are not able to understand atonal music.

Scruton’s theory does not take into account cultural influences and subjective musical experiences. Because there exist very different musical cultures and practices (such as, for example, free jazz, Chinese Peking opera, and some contemporary electronic music), we cannot reduce the musical experience to one which is universally of a tonal structural type. This view does not account for the variety of ways in which music can be organized. Moreover, it leads to the idea that tonal organization is a universal matter, or the ‘right’ way of hearing music. This does not find support in real musical practice, in particular if we conduct a cross-cultural analysis. Finally, tonality and its structures are, in fact, largely arbitrary; that is, they were not ‘naturally’ determined. As noticed by McClary, “the tonal process serves to ensure rational, platonic order, as theorized by Rameau around that same time” (2000, p. 97): this indicates that it is not a natural system, but a rationally established one.

Although I agree that there may be some common perceptual abilities and processes that are not culturally influenced – for example, some Gestalt principles –, it does not mean that people listen to music in the same way. These principles are universals only at a very general level, and they are highly influenced by one’s musical culture, which determines their way of listening to music. For example, a Western listener may hear Balinese music very differently from the way people of that culture listen to it. Moreover, the musical knowledge and the meaning a piece of music acquires within a musical culture shape a listener’s perception of music. This shows that they do not naturally hear music in a particular way: instead they learn how to hear it in a particular way.

In the article *Music and Emotion: Distinctions and Uncertainties* (2001), Meyer presents an updated version of his view on music and emotion, in which he supports the idea that understanding and responding to different styles of music depend most upon learning. He distinguishes between native processes, which are cognitively constrained and

---

103 This idea finds support in the thought of theorists like Rameau (1722), Pythagoras and his theory of the ‘harmony of the spheres; Zarlino (1558, 1571).
essentially universal, and syntactic processes, which involve learning the norms of a particular musical style.

An interesting concept showing that structural hearing is influenced by culture and musical knowledge is that of the ‘lexicon’. Bigand relates this concept to a tendency to abstract bits of musical knowledge and store them in long-term memory in the form of schema. “In every musical culture, rhythmic, melodic and rhythmic-melodic configurations recur frequently enough for listeners to store them in long-term memory in the form of a lexicon of schemas and proto-typical forms. Some experimental research has revealed the existence of a lexicon in Western musical culture” (Bigand, 1993, p. 242). The musical lexicon is thus related to the listener’s musical experiences and knowledge. It is important for the recognition of familiar music and it also influences listening to unfamiliar musical events because it determines how listeners build relationships between sounds. During her musical exposure, a listener acquires musical knowledge in the form of structures and schema that she then uses to organize music which she listens to for the first time. On this basis, she constitutes her musical experience as having a particular meaning.

Huron’s article Towards a Theory of Timbre (2002) is helpful in explaining the relevance of learning in music perception. It shows that tonality is not the only way to organize sounds through the idea of statistical learning. Listeners learn how to organize sounds in relation to the most recurrent structures and events occurring in the acoustic environment to which they are exposed. This then influences their notion of the probability that such events will occur again. On the other side, Huron’s theory might explain why, in most cases, listeners organize music according to tonality: this happens because it represents, in a culture, the most recurrent structure and, thus, the most probable.

Analysing tonality and the various feelings evoked by different scale degrees in listeners experienced with Western music, Huron addresses the question: “can we account for the feelings listeners experience by understanding the statistical patterns evident in the

\[104\] Many native processes are manifestations of the Gestalt principles of pattern perception.

\[105\] Huron analyses tonality form a psychological perspective, from which tonality might be defined as “a system of pitch-related expectations, where the various scale tones acquire distinctive qualia as artifacts of learned statistical relationships” (Huron, 2006, p. 174).
music itself?” (2006, p. 146). Referring to the aesthetic experience, Huron replies that, generally, we cannot. This is in line with an observation he previously made in analysing an experiment presented by Saffran, Aslin and Newport in the article *Statistical Learning by 8-Month-old Infants* (1996). These researchers were able to show that “what listeners heard as a figure during the experiment had nothing to do with the structure of the figures themselves, but related only to their simple probability of occurrence. The internal structure of the presented musical figures had no effect on the perception of grouping. Listeners based their perception on the statistical properties of various tone sequences. More precisely, they learned contingent frequencies. The case reported by Huron well exemplifies these results:

A simple linguistic analogy might help to clarify the results. Suppose you heard a long sequence of repeated syllables: abababababa . . . How would you know whether you were supposed to hear ab, ab, ab, ab, . . . or ba, ba, ba, ba . . . ? In effect, Saffran trained two groups of listeners, one to hear the sequence as ab, ab, ab, and the other to hear the sequence as ba, ba, ba. (In fact, in an earlier experiment, Saffran, Aslin, and Newport had done exactly this for spoken nonsense syllable.) For each item in the test phase, one group of listeners heard as a figure what the other group heard as a nonfigure and vice versa (2006, p.69).

The feelings that are experienced by a subject might be related to some structural features, but we cannot explain such feelings only on the basis of the actual composed structure of the music. The ability to detect structural changes, for example, might be related to other musical and experiential factors, such as learning and context. It follows that the perceived structure and the one that has been established while composing music do not coincide. This gives space to a subjective experience, since the characterization of a musical style depends on a person’s listening experience. Following this idea, the man’s perception of the musical events that he encounters in the metro is determined by his past experiences and exposure to music, which might be different from the actual organization of the music he hears.

*The process of similarity.*
The cognitive principle of similarity is particularly important in music perception. I argue for the idea that it is a universal mental mechanism which contributes to the listener’s structural organization of every kind of music. An example comes from the analysis of the work *The Angel of Death* by McAdams and Battier on the music of Roger Reynolds (2005). Among other goals, their study aims at exploring how listeners perceive similarities between original musical materials and variations or transformations of those materials during listening to a piece of contemporary music. In studying the listeners’ perception of Reynolds’ work through the creation of an infinite variety of transformations (often based on the manipulation of pitches and durations) and the presentation of grouping tasks (listeners were asked to group excerpts together according to their musical similarity), the authors found that the perception of similarity between different musical patterns determines their identification and recognition. They noticed that listeners recognize the recurrence of musical patterns according to their level of structural and surface resemblance. In other words, the detection of similarity is based on the perception of cues located at the surface as well as at higher hierarchical levels of musical structure.

This is interesting for the investigation of how listeners process music which is different from the Western tonal classical tradition and because it shows that people *learn how* to perceive music. McAdams and coll. (2004) in fact observed that listeners evolve a sense of familiarity with music while hearing it. The study also indicates that not only the hierarchical structure, but also surface features of music contributes to the perception of similarity. For example, the identification and subsequent recognition of musical excerpts can occur in relation to timbral similarity processing. In other experiments using Reynolds’ work, listeners were able to recognize a manipulated musical excerpt on the basis of timbral similarities. In this case, the structure which participants perceived during their recognition might not have been the same as the notated structure. I will return to this topic in the seventh chapter.

*Empirical studies supporting this perspective.*
In his article *Concatenationism, Architectonicism, and the Appreciation of Music* (2006), Levinson makes reference to some studies in experimental psychology that demonstrate the minor role of large-scale structure awareness in the appreciation of music. “These studies suggest that either such form is hardly registered by even experienced listeners, on any level, or else that to the extent it is registered, it makes little contribution to the sense that such listeners have of the coherence, expressiveness, or other valuable properties of the music being listen to” (2006, p. 511).

Tillman and Bigand, for example, in their work *The Relative Importance of Local and Global Structures in Music Perception* (2004), underline the influence of learning in the processing of tonal structures at the perceptual level (listeners become sensitive to the regularities of the environment) and share Levinson’s view for which “perception of music essentially consists in the experience of locally coherent musical moments” (2004, p. 219). Making reference to experiments in which global structures are manipulated (for example, modifications of cadences whose functions depended on the overall global context of the piece), Tillman and Bigand underline the fact that the results clearly question the perceptual impact of musical structures for the listeners. Perception seems to take place through a succession of local structures, and this leaves space for a more subjective structural organization.

Taking into account experiments about the perception of tonal closure,106 Elizabeth Marvin and Aleck Brinkman (1999) conclude that large-scale tonic closure has no effect on music’s perceived expressiveness. They observe that many theories proposed within the framework of Western tonal music analysis “assume that musical works will close in the tonic key, even when closure is preceded by tonal motion to other keys. It is unclear, however, whether this music-structural principle is actually perceived by most listeners” (1999, p. 390). Marvin and Brinkman assert that theories of musical structure need not be theories of musical perception. As they concluded from their experiments, “…participants seemed completely unaware that the large-scale musical structure of these works had been violated by the rearrangement of phrases, judging by their responses to our “stylistic

---

106 Tonal closure is the conclusion of a musical composition on the tonic harmony of the key in which it began.
conformance” question” (1999, p. 406). Thus, it is difficult to claim that the perceived structure always correspond to the notated one: if this were the case, listeners would have detected structural changes, even if they did not change the piece’s expressiveness in a substantial way.

A similar study has been presented by Nicholas Cook (1987). He questioned theories like those of Schenker, Meyer, Lerdahl and Jackendoff, which explained the aesthetic effect of music in terms of hierarchical relationships between large-scale tonal structure (like tonal closure) and the local events of a given composition. Through his experiments, Cook showed that a theory of musical structure does not have to be a theory of perception and that musical structure may be of a conceptual or perceptual nature. “Viewed in this light, theories that explain the organization of Classical and Romantic compositions in terms of large-scale tonal structure may not correspond in any direct manner to the perception of such music, but they may still be of value in revealing something of the manner in which composers of the tonal period conceived their music” (Cook, 1987, p. 204). It is therefore necessary to distinguish the conceptual from the perceptual nature of a musical structure because the first is not always perceived the way it is written. Additionally, the conceptual nature of musical structure does not necessarily have a direct influence on the appreciation of music. This also confirms the idea that listeners (in particular non-musicians) pay more attention to local structures than global ones. The large-scale tonal structure may not in itself be perceptible, but it plays an important role as a means of compositional organization. Presenting short pieces of music from three to six minutes in length whose closing sections were transposed into a different key than their opening sections107 (thus undermining tonal closure) Cook found that the key changes where not detected by musical subjects, and had no effect on experience coherence, completeness, expressiveness or pleasurability. The direct influence of tonal closure on listeners’ aesthetic responses were relatively weak and restricted to short time spans.

---

107 Tonal music generally begins and ends in the same key.
Heidi Gotlieb and Vladimir Konečni (1985) have also observed that a lot of theorists believe structural modification will change a piece’s aesthetic effect. To challenge this notion they have conducted experiments in order to question the importance of the structure. The conclusions drawn from such experiments indicate that variations and modifications to an original piece of music have only a minimal effect on subjects’ enjoyment of it. In another experiment Konečni (1984) showed that drastic global structural alterations of Beethoven’s string quartets and piano sonatas resulted in no measurable advantage in enjoyment or comprehension over the original versions, according to participant preference and familiarity ratings. In an experimental study that he conducted with Mitchell Karno (1992), the results showed that the ability of even musically trained individuals to perceive structural changes is highly suspect.

These results are contrary to a lot of other empirical studies. To make a comparison, we have to distinguish between studies which analyse detection tasks about structural changes from studies about the aesthetic response listeners have in correspondence to structural changes. These refer to two different kinds of perceptual experiences, but neither implies that the perception of the ‘correct’ notated structure have a determinant influence on the constitution of musical experiences.

A good example showing that the perceived structure might be different from the notated/established one is in the case of meter perception. The perceived metrical structure is determined by an organization of beats which is sensed on the basis of their accents. Theodore Zanto et al. (2006) describe meter perception as the perception of regularity in accent or stress timing, giving rise to a sense of strong and weak beats. Emilios Cambouropoulos claims that “metre is not simply a mental artefact induced from the music but actually has an autonomous psychological existence that is developed within a cultural context and influences actively the way music is performed/perceived” (1997, p. 289). This allows us to consider meter as a musical element which is not strictly related to the compositional structure, but open to different subjective interpretations.

*The case of unfamiliar music.*
As we have noted both in discussing Huron’s theory and according to the enhanced concatenationsim presented by Levinson in the second part of his work *Music in the Moment*, listeners learn how to perceive music structurally. Tonality does not reflect a natural way of hearing music, but it is learned, consciously or unconsciously, through exposure. What happens when a ‘tonal’ listener is exposed to music that she has never heard before or which is not tonally structured, such as medieval monodic music or some contemporary electronic composition? She probably does not have the ability to organize sounds according to the notated/established structure (if any) of that particular music because she is not familiar with it. In Scruton’s opinion, the consequence is that she does not understand music and she cannot appreciate it.

However, I argue that in this case she is still able to create relationships between sounds, which allows for a basic musical understanding of sound events. A sequence of sounds is constituted in music thanks to a structural organization which is based on the subject’s musical experiences. In the case of unfamiliar music, even if the perceived structure differs from the notated one, the listener is able to constitute musical experiences. The way she hears music determines her appreciation and perceptual comprehension of music. For example, she can appreciate and recognize Schöenberg’s *6 Kleine Klavierstucke op. 19 (1911)*\(^{108}\) or an Indian raga as musically meaningful even if she knows nothing about their structure and perceives them differently from culturally accustomed listeners. These might be considered examples of incorrect musical understandings, but they are nevertheless meaningful musical experiences. It will be interesting to discuss the role of expectation and anticipation in this process.

When listeners are unable to make schematic distinctions and they judge the sound events that they hear by saying ‘It all sounds the same to me’, it is because – in Huron’s

\(^{108}\) Arnold Schöenberg developed the twelve-tone technique, which has been a widely influential compositional method involving the manipulation of an ordered series of all twelve notes in the chromatic scale. He was the first modern composer to embrace ways of developing motifs without resorting to the dominance of a centralized melodic idea. Due to the great difference of this compositional practice with respect to tonal classical music, the complex structures of his compositions are often unknown to ‘tonal’ listeners.
opinion – they have not acquired the schematic boundaries suitable to distinguish between the sound events. Huron suggests that “listeners may simply fail to gain sufficient exposure to bring about the creation of the new schema. Such failures are commonplace when listening to the music of an unfamiliar culture” (2006, p. 215). Since it is possible to learn new schemas through mere exposure, listeners will eventually find new ways to organize sounds. In any case, this does not mean that they will learn to perceive musical structures in exactly the same way in which they have been established. They may create a subjective way of hearing these structures which allows for the most appreciation and enjoyment of the music.

Thinking in sound.

Some theories (both musical and psychological) have compared music with language. In these cases, music is seen as a symbolic system referring to a semantic content that is not fixed.\(^{109}\) Music listening is considered as a ‘thinking in sound’ process, characterized by the use of grammatical rules that allow the comprehension of music.\(^{110}\) Listeners correctly understand music when, following the syntactic system of the music to which they are exposed, they are able to grasp the structure that establishes the identity of a musical work (the one which is usually notated in the score). The idea that the work has a structure in itself and that it is an identity condition for both recognition and comprehension poses the problem of what a musical structure exactly is, and to which ontological category a musical work belongs. If the notated structure is a necessary identity condition of a musical work, then the ability to grasp it correctly is fundamental for the comprehension of music and its recognition.

We may consider the identity of a musical phenomenon as something that contains both an abstract and a physical aspect. It has an objective existence in the real world because it is constituted by physical sounds. But it also has an abstract existence because of

\(^{109}\) In the case of emotion, for example, in Langer (1942) and Goodman’s (1968) opinion music is symbolic precisely because it is experienced as resembling or exemplifying what it denotes. For an overview see also Davies (2001a); Cook & Dibben (2001).

\(^{110}\) As we have seen, this view is supported by Lerdahl & Jackendoff, 1983a.
its shape, which is determined by relationships that are perceptually created between sounds. Music is organized at the physical level as an ordered sequence of sounds, but also at the abstract level as a structural form: on the basis of the perceiver-environment relationships, this allows music to take life as a particular human experience. These two aspects – that is, the abstract and the physical – are indissoluble and fundamental to musical experiences.

If we reflect on the experience of human beings in the world, we see that each individual experiences different phenomena, each of them having its own structure, form and identity. Groups of people share the same environment and they are in constant interaction with it. They can all share common interaction with a particular phenomenon, but each of them has a different experience of it. They live in different ‘points of view’ in the world and they have different experiences of the same musical phenomenon. This is due to the fact that individuals develop their abilities in different ways, and they are subjected to different experiences and cultural influences. In this way, they develop subjective perceptual processes and schema, which form the basis for the creation of their own musical experience. It follows that listeners may structurally organize the same piece of music in different ways because of the unique ways of hearing music that they have developed. This also means that the perceived structure does not necessarily coincide with the notated/established one, without implying that music cannot be appreciated or perceptually understood. Thus, there is an irreducible subjective aspect that determines different ways of listening to music, even if people learn the regularities of the musical system which is dominant in their culture and hear music structurally. Thinking in sound, listeners constitute subjective meanings in relation to the same acoustic event.

111 We can find good examples of this aspect of the musical experience in the article Perceptual Structures for Tonal Music by Carol Krumhansl (1983). The author discusses some investigations into the psychological representation of pitch relationships in the case of tonal music. In particular, she reports evidence for the internalization of tonal structure at three levels of organization: musical tones, chords and keys. The results of the presented experiments suggest that the listener relates the sounded elements to an abstract internal representation of the structural regularities underlying tonal music.
Final reflections.

The idea that I have tried to present in this chapter is that listeners perceive music in a structural way, but they do not necessarily need to grasp the notated/established structure of a musical piece in order to constitute musical experiences and to understand the sound environment as music or in musical forms at the perceptual level (as the man in the metro does). This not only allows for the appreciation and a kind of perceptual comprehension of music having unfamiliar structures, but it leaves space for a subjective way of hearing sounds. In Klaus Scheree and Marcel Zenter’s terminology (2001), it allows the development of the musical expertise.

The problem of perceived versus notated structure is related to the question about the recognition of two musical works as the same. In philosophical terms, when we say that A is the same as B, we can ask ‘the same what as B’? The recognition of two musical events as the same is not made only on the basis of some structural features or the fact that they share specific structural elements. It depends on other factors as well, such as: superficial aspects, contextual elements and past experiences with those musical events. Therefore, it is very hard to define what it is that determines our claim that an event A is the same as event B. Moreover, I argued that the perceived structure is not necessarily the same as the notated one because listeners develop different ways of hearing music and different ways of organizing sounds structurally. This not only creates some problems for the idea that the structure is an identity condition of the musical work, but also for the causal account theory of perceptual musical experience (the same tonal system is characterized by cause-effect relationships between sounds). As Cook notes, perceiving music is different from conceiving it.

Cohen observes that “quite apart from the complexity of music itself, well documented in the centuries of scholarly writing about music, additional difficulties beset the description of the mental representation of musical structure. Individuals differ widely in their musical abilities, depending on age, experience, attention, hearing mechanism, innate intellectual propensities, etc” (1982, p. 17). As noted previously in the chapter, the
constitution of musical experiences is not limited at the perceptual level to the perception of the ‘correct’ notated structure. One problem concerns the extent to which it is not limited in order to guarantee a musical social communication and the recognition of a repeated piece of music as the same. The elements (structural or otherwise) that determine the identity of a musical work must be grasped at some level by more listeners in order for musical communication to be realized. Such a level is difficult to establish, particularly in the case of music perception that does not involve a conscious reflection on what is heard.

Unfortunately, it is very difficult to empirically investigate the subjective musical experience and perceived structure. As Adorno writes, “to make scientifically sure of the subjective content of a musical experience, beyond superficial indices, is an all but prohibitively difficult task. Experiments may tell us about degrees of the intensity of the reaction; they will hardly reach its quality” (1976, p. 4). For this reason, empirical studies try to objectify the perceptual experience, but they are not able to give a complete account of the perceptual structural experience. The risk is that in attempting to objectify musical experience, they may explain it only according to the way music has been notated and on the basis of the syntactical rules that characterize a particular musical culture.

Some theorists reject the cognitivist point of view because they do not agree with the idea of ‘stimuli’ in perception. This implies that the perceptual experience is limited to the context of the laboratories in which experimental studies are made, and thus they do not consider the real context in which music is subjectively perceived. Only recently have researchers begun to develop studies in real time contexts.112

Another aspect that some theorists like Clarke (who supports an ecological approach to music perception and refuses a structuralist music theory approach because it is score-based and abstract) have criticized about the cognitivist point of view is the association made between sound and representation. This brings up the question of what representing sounds in structural form means. An interesting consideration is that the association of sound and representation could be viewed as the outcome of a collective training. Representational systems have particular properties that go beyond their purely perceptual

112 For example, McAdams, Reynolds & Battier (2005).
attributes, and understanding music in this case requires the knowledge of a particular language or, in Lerdahl and Jackendoff’s terms, a musical grammar. From this perspective, the representation of musical structures depends on specific musical knowledge and experiences that influence the perceptual organization of music. This draws upon the idea that listeners understand a piece of music only if they organize and represent it according to its notated/intended structure.

However, the subjective aspect of the structural organization of music leaves room for different interpretation and organization of the sounds constituting a musical composition. The perceived structure is a version of the notated one, which contains an irreducible subjective aspect. As Brough writes in the preface of his translation of Husserl’s *On the Phenomenology of the Consciousness of Internal Time (1893-1917)*, “the contents and apprehensions that constitute the perception of the immanent temporal object are not identical with the object itself” (Brough, 1991, p. XLVI). This quote is useful in highlighting the subjective character of the man’s musical experience in the metro station. The musical structure that he perceives on the basis of an unconscious organization of sounds is determined by his knowledge and experiences, but also by his context and situation. Such structure influences his perceptual awareness of music, as well as the meaning that his musical experience acquires in that context. For example, the unexpected character of his experience makes him perceive music differently from a man who goes to a concert: different concepts, intentions, contexts and expectations are involved.

**The role of expectation.**

What is the role of expectation in the structural organization of music at the perceptual level? As we have seen in the previous chapters, the creation of expectations is based on experience and some perceptual abilities. I suggested that they allow for different perceptual understandings of a musical event. Because the expectation process underlines the subjective aspect of the musical experience as well as the relevance of learning on the
constitution of musical structures, it may also support the distinction between perceived and notated structure. Different structural aspects influence the creation of expectations, which are highly determined by memory of past experiences. Juslin and Västfjäll, who consider musical expectancy as a syntactical processing, claim that “musical expectancy refers to those expectancies that involve syntactical relationships between different parts of the musical structure” (2008, p. 568). The influence of previous encounters with the same musical system is important in this respect: like language, music consists of perceptually discrete elements, organized into hierarchically structured sequences according to the rules of a particular musical system. Thus, it is a common view among music theorists that most musical styles are, in principle, describable by a grammar (Lerdahl & Jackendoff, 1983a). It is only through the perception of this syntax that the relevant musical expectations arise. These expectations are based on the listener’s previous experiences with the same musical style (Carlsen & Krumhansl, 1999).

Emotional reactions to music are induced when the listener’s musical expectations are somehow disrupted, for instance, by a new or unprepared harmony. Moreover, “musical expectancy reflects learned schemata about specific styles of music that differ from one culture to another and that make listeners from different cultures react differently to the same piece of music” (Juslin & Västfjäll, 2008, p. 569). This describes the influence of cultural exposure and the relevance of learning for the creation of expectations. In this sense, Altenmüller (2004) has presented an interesting concept: that of a ‘learning biography’ that each listener develops at the neural level during her exposure to music.\textsuperscript{113} The basic idea is that cortical activation during music processing reflects personal experiences accumulated over time. As the author writes, “listening to music, learning to play an instrument, formal instruction, and professional training result in multiple, in many instances multisensory, representations of music, which seem to be partly interchangeable and rapidly adaptive” (2004, p. 346). Through experience individuals thus develop quickly adaptive brain substrates, including widely distributed neuronal networks in both hemispheres.

\textsuperscript{113} See also Tillmann, Bharucha & Bigand, 2003.
In the case of expectation theorists sometimes make the error of studying a listener’s response to music by examining the music itself (an easy error in the case of tonal music which is built on expectations) and without reference to the real context in which music is perceived. They attempt to objectively discuss a subjective content and musical experience as Meyer did in his book *Emotion and Meaning in Music* (1956) by examining the affective response to music which is, according to the author, highly influenced by the structure of the notated composition.

Taking into account the analysis of tonality made by Huron and the process of expectation, we see that Western-enculturated listeners are able to create correct expectations for some aspects of the scale-degree distributions (like pitch proximity expectations). This does not mean that they always grasp the musical structure as outlined by the composer and the music theorist. Non-musician listeners, for example, do not have precise comprehension of notated musical patterns or structures. They develop different musical understandings and meaningful experiences in relation to other factors, like context and intentions, and the expectations such factors determine.

As we have seen, expectations are usually related to the comprehension of the organizational features of music, but they do not always accurately reflect the way musical sounds are actually organized. Firstly, they are influenced by the listeners’ musical knowledge and experiences, which do not always involve the music they are actually listening to. Secondly, listeners are not perfect in their musical expectations and, when they anticipate sounds they often rely on heuristic knowledge. It follows that listeners do not always grasp the correct organization of musical sounds. Finally, the creation of expectations depends upon different factors including cognitive and perceptual abilities, acquired knowledge and structures, subjective experience, context and situation.

These points are clear in the case of the man in the metro. For example, he might perceive a well known piece of music in a very different way from his past experiences with it and the way it is notationally structured because of the complex auditory environment of the metro station. Other sounds and the unexpected character of his musical encounter makes him organize sounds in different or new ways at the perceptual level. This
does not imply an ‘incorrect’ understanding, but it determines a different musical experience which may acquire a different meaning. As we will see, expectation and anticipation play a great role in this phenomenon.
LISTENING IN THE MOMENT AND THE ROLE OF ANTICIPATION VERSUS EXPECTATION

Introduction

An aspect of every perceptual act which is fundamental to the immediate awareness of the various meanings that events occurring in one’s environment acquire is its structural unfolding in time. This structural unfolding plays a crucial role in the development of what has been called ‘listening in the moment’ and in the basic perceptual comprehension of both familiar and unfamiliar musical events. The main idea is that listeners constitute their musical experiences in the moment of their perceptual acts in the living world due to the development of structural relationships between sounds. The temporal structure of human consciousness, which is characterized by a complex intentionality, may help us to understand what listeners mentally do to hear certain sounds as music and how they are aware of these perceptual experiences. In particular, it could clarify how expectations are built during the musical experience and show their influence on the perceptual comprehension of music.

In order to discuss this hypothesis, I will outline several studies and theories dealing with the specific role of expectation and the factors which distinguish it from protention and memory. I will relate these ideas to the distinction between expectation and anticipation which was introduced in the fourth chapter in order to better understand the influence of expectation on the constitution of subjective musical experiences.
The temporal structure of human consciousness.

Background and perspectives.

In this paragraph I analyse the temporal structure of human consciousness from a phenomenological perspective. In particular, I take into account some secondary literature addressing Husserl’s essay on the awareness of inner temporality and its structure (1893-1917).

One of the phenomenologist’s aims is to uncover the essential structures of the intentional consciousness. Husserl, for example, sought to explain the \textit{a priori} structures of the temporal experience as such, through an approach which was radically different from that of the empirical sciences. While the latter try to explain the human consciousness objectively, phenomenology accounts for the subjective aspect of the temporal experience. Husserl, in fact, investigated the perceptual subjective experience of time and temporal objects.

In the present study, I do not focus on the experience of time as a form of temporal awareness, and I do not discuss the features of temporal objects. The interest is not in the consciousness of time (as, for example, the awareness that the song I am hearing is ‘five minutes long’), but on the temporal movement through which listeners are aware of music \textit{during} the perceptual act of hearing. I follow Husserl’s account of the nature and structure of the awareness that characterizes the absolute flow.\footnote{The absolute flow constitutes the internal time of all intending acts and, through them, the transcendent time of all external objects. In addition, it also constitutes itself.} For this reason, the focus of this discussion will be about the \textit{structure} of our experience of a temporal object – that is, the formal structure of the temporal movement of human consciousness.

As Dermot Moran observes, “we are not conscious of time as such, but rather of objects \textit{in} time. We next have to distinguish temporal objects and their parts from the inner temporal structure of the \textit{Erlebnisse} that present these objects (the phases of the musical tone have to be distinguished from the temporal phases of the hearing experience, although
there is clearly a parallel between them)” (2005, p. 139). My analysis concerns this second aspect. The aim is to understand how human consciousness works in order to relate sounds in time and have perceptual musical experiences. Moreover, I shall try to explain the moment-to-moment musical listening and underline the role of expectation and anticipation through an analysis of the temporal phases that characterize temporal structure. This structure is fundamental to the musical experience because it allows for the perception of sounds as music and for the constitution of subjective musical experiences.

Features of the temporal structure.

We may start this discussion by observing, as Husserl did, that people in general perceive successive and enduring objects. In the case of music, this refers to the listener’s ability to perceive a sequence of sounds as a meaningful and unitary whole. This ability is what enables the man in the metro to perceive music; to understand a temporal sequence of sounds occurring in a complex acoustical environment as music. Such an ability to perceive enduring objects is therefore at the basis of the constitution of musical meanings.

A relevant issue concerns comprehension of the main features of such perception and the way it is realized. One reply to the problem requires analysis of the temporal structure of human consciousness and the moments that characterize its development. These moments include: retention, primal impression and protention, and they depend on one another in order to create the impression of a cohesive whole out of successive temporal events. Through a combination of such moments, listeners can have continual perceptual experience and remain aware of the sameness of a musical object. In this way, a sequence of sounds is perceived as a meaningful musical unit. To support this idea, I base my discussion on the literature which analyzes Husserl’s lecture on time-consciousness, which is a phenomenological-eidetic examination of how the continuity of consciousness and the consciousness of continuity are possible.
In analysing the structure of time consciousness, Husserl focuses on the present moment of the perceptual experience, which he explains by referring to its three characteristic time phases: retention, primal impression and protention. These characteristics are fundamental to what I have called ‘basic musical understanding’ and for the experience of music in the moment: they allow the creation of continuity and movement between sounds throughout the temporal duration of a musical segment. They also guarantee unity, or a perceived musical whole, from our temporal experience of sound events.

The first problem to take into account – prior to the analysis of these three time phases – concerns the distinction between ‘experience’ and ‘perception’, and their meaning in relation to the temporal structure. Retention, primal impression and protention are not forms of perception or acts of perception, but modes of impressional consciousness. Brough clearly explains this point:

Unlike perception or memory or reflection, which thematize their objects (I see the flying bird, I reflect on my act of seeing it), impressional consciousness might best be described as the nonthematizing awareness of what is immanent to consciousness, such as an act of content in its temporal extension. It is the implicit self-consciousness that always attends my conscious life. Husserl frequently calls the immanent act or content an “experience” [Erlebnis]. In the case of perception, a transcendent object – the bird, for example – is perceived through the act. The act itself, the “Erlebnis” is not perceived; and yet I am aware of it: it is “experienced” [erlebet] or “intended” [bewusst], Husserl says. Similarly, in an act of secondary memory I recall thematically a past object or event; I do not recall my present act of remembering, nor do I “perceive” it (unless I institute a specific reflective act directed towards it). I do, however, “experience” or “intend” the act as the mental content I am presently living through. This experiencing of the act as an immanent object extended in immanent time is precisely the accomplishment of the absolute consciousness –an accomplishment that the absolute flow brings about through the retention, primal impression, and protention that make up each of its phases (1991, p. XLIX-L).

These modes of impressional consciousness, or temporal phases of the absolute flow, are experienced and realized as part of the immanent act of consciousness. We are not conscious of the processes that characterize perception, but we are aware of our perceptions through these three phases. As John Brough notes, “these are no longer taken to be names

---

115 Husserl conducted his analysis of time consciousness between 1909 and 1911, when he clearly presented the different levels of time-consciousness, the different intentional correlates corresponding to them, and the place of perception in the larger schema of temporal awareness.
for moments belonging to a perceptual act; they are rather moments of the ultimate level of consciousness through which one is aware of the perceptual act – and of any other act or content – as an immanent temporal object. Perception now becomes simply one more act of consciousness constituted, along with memory, expectation, judgment, phantasy, and so on, as an immanent temporal object through the absolute flow” (1991, p. XLIX). The relationship of the three phases to the perceptual act is also clear in Moran’s quotation: “a retention is a part of a perceptual awareness; it is a ‘just past’ that is still there in a reduced or modified sense. It still has a kind of ‘impressionality’. Similarly, protention is not yet the fully-fledged conscious act of anticipation, but a structural component of any *Erlebnis*” (Moran, 2005, p. 142).

As I have previously underlined, the temporal structure is related to the perception of sameness: this allows for the perception of a sequence of sounds as belonging to the same musical event, even within a complex auditory environment, such as the case of the man in the metro. The perceptual awareness of a melody across an interval of time, for example, necessarily involves a continual awareness of the succession. As Izchak Miller observes, “at any given moment we retain the content of our earlier perceptual act. Thanks to our retentions we can experience the current (purpoted) object of our act as being the same (purpoted) object experienced by us earlier” (1984, p. 81).

In referring to the modifications that a form bears during its temporal development, Husserl claims that “each mental process . . . is followed . . . by a “retentional” consciousness . . . by virtue of which the primitive mode, “given at present”, goes over, in a continuous synthesis, into the modified form, the Same that “just now” was . . . .This undergoing of continuous retentional modification is the essential initial part of the constitution [the consciousness] of an identical object, one that, in the broadest sense, persists” (as cited in Miller, 1984, p. 72). Thanks to these modes of appearance, a listener is aware of an object as enduring, individual and concrete. She perceives a whole or sequence of sounds as a unity enduring through time. In the case of music, these modes of appearance allow the perception of, in Husserl’s words, a “temporally extended melody, consisting of many successive tones, through a single perceptual phase that enjoys its
moment of actuality and then is gone” (as cited in Brough, 1991, p. XXXIII). This also
indicates that conscious perceptions, or acts of consciousness, are themselves processes.
Consciousness may thus be seen as a perceptual Heraclitean flux\(^{116}\) and human beings are
aware of objects because of the temporal formal structure.

Focusing on the present moment of the perceptual act, we can see that retention,
primal-impression and protention are strictly related. Although we can distinguish them,
they are dependent modes of a momentary phase of consciousness, of the act of perceiving
an object and of a whole of temporal appearance or temporal giveness. They refer to one
another and, in this way, time-consciousness does not just refer to a punctual now, but a
now which includes a horizon of past and future. In his article *Music as a Temporal Form*,
Joan Stambaugh writes: “this is the very essence of musical motion: the constant creation of
a future and a past in the actual present moment, in *each* present moment” (1964, p. 276).
The present is principally a horizontal form which persists through continuous changes of
content. It corresponds to the duration of the event or object of which the subject is aware
during the act. The present is thus the duration of a perceived enduring event that involves
retention, primal-impression and protention. In Sokolowski’s words, “the term *the living
present* signifies the full immediate experience of temporality that we have at any instant.
The living present is the temporal whole at any instant. This living present, as the whole, is
composed of three moments: *primal impression, retention, and protention*. These three
abstract parts, these three moments, are inseparable” (2000, p. 136).

\(^{116}\) The Heraclitean theory of perpetual flux and universal transformation states that everything is constantly
changing. It is characterized by Heraclitus’ proverbial utterance: ‘All things are flowing’. Before Heraclitus,
the Greeks focused on the essence of nature and being, and the world was considered fairly static. Everything
was firmly embedded into an indivisible universe and the common principles of nature were perceived as
unchangeable. In contrast, Heraclitus said: ‘You cannot step into the same river twice, for fresh waters are
ever flowing in upon you’. This simple sentence expresses the gist of his philosophy, meaning that the river is
not actually the same at two different points in time.
He looked at everything as being in the state of permanent flux and, hence, reality being merely a succession
of transitory states. His ideas regarded the notion that nothing is the same now as it was before, and thus
nothing that is now will be the same tomorrow. With this he planted the idea of impermanence into Greek
thought, and indeed, after Heraclitus Greek philosophy was not the same anymore.
Heraclitus states that every object is subject to change and is always undergoing some kind of change or other
through the concept of the ‘unity of opposites’. The point is not that *everything* is changing, but that the fact
that *some* things change makes possible the continued existence of *other* things. Perhaps more generally, the
change in elements or constituents supports the constancy of higher-level structures. The message is that
rivers can stay the same over time even though, or indeed because, the waters change.
In the case of music, the temporal structure is fundamental for the constitution of relationships between sounds and their perceptual realization into a unitary whole. Hearing a melody, for example, essentially involves hearing its constituent tones in succession; the tones are constituent parts of a single process. Each tone of the melody expresses a part-whole attribute which determines the whole utterance in question. Each tone is not perceived as an atomic, single impression, but in relation to a whole that is subject to modification and characterized by continuity. Putting sounds in relation to one another, the listener perceives changes in a process that is essentially dynamic. This indicates that the structure of retention and protention is not rigidly fixed, but always in process.

Husserl’s description of the manner in which a melody is experienced well contains all these points – duration of perception, modification, continuity and awareness of continuity, parts-whole relationships and temporal movement between the three phases of the temporal structure – and I report it entirely through Lawrence Ferrara’s words:

A melody is a temporal object. It is comprised of different tones that constitute a section or the overall melody itself. One must wait for the very last tone of a section or of the entire melody before he can fully comprehend its overall effect. Each tone endures for more than its actual sound. In other words, a quarter note may actually sound for only one beat (perhaps less than one second of chronological time) but it may endure in consciousness for the entire melody (in musical time). We hear the tone “now”. As that sound moves into the past a new tone sounds. However, the original (past) tone is still enduring (or sounding in our inner ear) while other tones of the melody sound in their respective “nows”. We continually modify the original tone as the rest of the melody continues to be played. Each tone in the melody contains this temporal density. Each tone is both now and retained (undergoing continuous modification) in our consciousness. While hearing each tone of a melody and continually modifying the meaning of past tones, we anticipate or “protend” where the melody may lead. The very ground for an object becoming a temporal object is based, according to Husserl, on this “duration of perception”. Temporal duration is characterized by the thickened quality of tones understood not in flat chronological time but in the enduring, ever overlapping and simultaneous “hearing” of past, present and future in musical time (1991, p. 148-149).

I propose to further analyse the temporal structure – its phases and their relations – because it plays a crucial role in the constitution of meaning and meaningful musical experiences. It is because of the temporal structure that various components of the perceptual experience (i.e., sounds and noises) are intentionally interconnected in order to

---

117 See Husserl’s lectures On the Phenomenology of the Consciousness of Internal Time (1893-1917).
constitute meaning. It is a fundamental condition for the organization of sounds into musical forms, which are understood as unities that are progressively built in time.

Constitution is a term used by authors such as Thomas Clifton “to describe a process by which the person orients himself toward a particular object which assumes significance for him in a certain way” (1983, p. vii). In the case of the man in the metro, this means that he intentionally organizes the sound environment in a meaningful way. A sequence of sounds are constituted into music and organized to form a subjective musical experience. All this happens in ‘time-consciousness’ thanks to the temporal structure. Joseph Smith describes ‘time-consciousness’ as “the primordial place for the constitution of musical identity and unity, as well as the source of the connective forms of coexistence and succession of objects in consciousness. Time-consciousness is, in fact, the origin of what we know as “form”, for in it entities shape themselves in passive synthesis for consciousness” (1979, p. 112-113).

A relevant feature of this process of constitution is its synthetic character, to which the three phases of the temporal structure are related. Such synthetic character manifests itself in the intentional framework of consciousness and, in the case of retention, primal impression and protention, it describes an identity synthesis that is achieved in acts of perception. As Moran observes, “synthesis is ‘a mode of combination exclusively peculiar to consciousness’. It is through synthesis that conscious experiences connect together into a unity, and that an identical object is grasped in the manifold of appearances. […] Identification is the fundamental form of synthesis, and the most basic form of this synthesis occurs in our internal consciousness of time” (2005, p. 152).

Another important concept related to comprehension of the temporal structure is that of ‘horizon’. This indicates that the listener’s present experience of an object is temporally extended beyond the present moment of experiencing it (i.e., it also includes what is not sensorially experienced in consciousness). In this way, it extends the size of the present to the past and the future. The identity of the present is established by what the past and the future ‘see’ of it, but the possibilities of relationships between them never cover the past and the future completely.
As Sokolowski writes,

> [t]o explain how we experience time, Husserl says that the now-moment is not given as an unextended point in our consciousness dividing past from future; rather the present is given together with its horizons. It is given with retention of now-moments that have just elapsed and with protention into imminent “nows”. Our awareness of the present includes immediate, direct awareness of retentional and protentional phases. These are known as “just having occurred” and “just about to occur” (1964a, p. 85).

I do not agree with the idea that we are aware of retentional and protentional phases because they are experienced but not perceived. However, I do feel that the above Sokolowski quotation is helpful in clarifying the concept of ‘horizon’.

Finally, I would like to discuss the concept of intentionality. As we have already seen, every act of consciousness must be said to be intentionally directed toward an object. The consciousness of time represents a complex form of intentionality. Husserl, in fact, explains the temporal perception of an object as having a three-fold intentionality. In this framework, the temporal phases of primal impression, retention and protention are identified with respect to partial intentions and intentional features of the mind, even if they are not acts of consciousness (as, for example, the act of perception).

A problem that may arise in analysing the interaction of intentionality with time is: what is the direction of time? What are its features? The directionality of time is important in order to understand the intentional movement of human consciousness and the role of expectation. The temporal structure is characterized by an intentional movement toward the future which is influenced by a relationship with the past (in the sense that, during perception of enduring objects, the present relation to the past lies in this future-oriented intentionality). Time, instead, is in itself ‘timeless’, it has not a point of departure, nor a point of end. It is symmetric in the sense that it has no direction. For the human experience, instead, time is asymmetric in the sense that, starting from the present point, it goes either in the direction of the future or in the opposite direction toward the past. This is a distinction that represents an intrinsic feature of human cognition: a tendency to ‘make order’ and give sense to experiences in the world. The distinction between past and future and the direction of time exist only in the subject’s experience, which is characterized by an
intentional action. This allows for the constitution of meaningful experiences, the knowledge and familiarity with the living world.

From this perspective, time is also irreversible and unidirectional: it always goes toward the future. The intentionality and future-directed action of temporal movement may be understood from a teleological perspective. Every act of consciousness is driven by an intentional movement and every element of its structure is a function of future-oriented movement. I will consider the idea that the elements of the temporal structure are functions of consciousness by discussing Gallagher’s critique of Husserl in the following sections.

The three phases of the living present.

In this paragraph I introduce a more detailed analysis of the three modes and phases of the temporal structure: retention, primal impression and protention, with particular attention to the first and the latter.

Retention.

Retention (first called ‘primary memory’ by Husserl) has been defined as the impressional consciousness of the just-past phase of the temporal object. It is the original consciousness of the past, in which what has elapsed is still present as the just-past, in the sense that its being past is something in the present. It is the immediate consciousness of the ‘just-having-been’ and it is responsible for being aware that a continuous group of tone-phases has sounded. Philosopher John. Brough writes: “Retention is one of the ways in which I experience something appearing as past in relation to the now” (1996, p. 9). In this way, different moments of the present experience are related into a unitary meaningful whole.

Retention is not a past moment as such; it does not re-present what has just elapsed in the way in which ordinary or ‘secondary’ memory would. It is a present consciousness (a retaining) of the past – given within the present. This allows that elapsed points of
experience are still present in consciousness. Retention is part of the experience of an extended object, but it is not perceived as such – thus, we are not aware of our retentional phases, even if they are at the origin of our awareness of persisting objects and perceptual acts. Retention refers to an intentional activity characterized by continuous modification of one’s orientation. It is not itself an act, but a dependent moment which is part of the triple intentionality. Its intentionality presents the past and gives the past originally. The following Brough quotation offers a good definition of retention:

> The retentional consciousness would arise in the following fashion. Assume the first tone (A) of a melody is apprehended as now, and then elapses. In the next phase of consciousness a new content replacing the old A will be present (generated there by a process which can only be described as magical), and the consciousness of A as just past will be constituted when this content, really present in the actual phase of consciousness, is animated by the past-apprehension. [...] This means that "the retantional tone is not a present tone but a tone ‘primarily remembered’ precisely in the now”. [...] (ZB, 31). (1989, p. 275-276).

What is retained is not static, but dynamic, still living, not yet finished, and caught up in continuous change.

A problem that arises with regard to the concept of retention is: how far does retention extend? This may be addressed by taking into account the difference between retention and secondary memory. Secondary memory is a ‘richer’ form of consciousness. It is not a perceptual act, but it re-presents an object. It assumes that a sense of the past has already been established in primary memory (retention), which is the primitive or first form of the constitution of the past. Secondary memory is itself an independent and intentional act with its own successive phases which re-presents its object. Whereas retention is presentative – that is, it contributes to the making-present of an experience – secondary memory is re-presentative. During the living present the object of consciousness exists in real time and retention contributes to its experience, while in memory a past experience of an object (or an object experienced in the past) is recalled and re-presented to the consciousness.

*Primal impression.*
In the case of hearing music, primal impression is responsible for one’s awareness of an instantaneous tone-phase as sounding-now. The now, or present, is a mode of appearance: the mode of appearance of that which is present. The ‘now’ is the source of the new, or consciousness’ aperture to the new. “The now, Husserl claims, is the point of reference for consciousness, the point of orientation in terms of which what appears as absent in our temporal experience organizes itself” (Brough, 1996, p. 8).

**Protention**

Protention (also called ‘primary expectation’) is, in musical terms, the consciousness of a future continuity of tone-phases. It is the primitive or first form of constitution of the future and represents the third moment of the intentionality of temporal structure. It is the immediate consciousness of the future which is present in all experiences. Clifton explains it in this way: “protention is the term for a future which we anticipate, and not merely await. Awaiting, like recollection, implies a disengagement from the present, whereas, experientially, the now which we perceive is coloured by the way we intend a future. Intending a future with respect to a given event means to attach significance to that event in proportion to the way the present and future are attached to, yet distinguishable from, each other” (1983, p. 62). In its relation to the horizon and the possibilities the horizontal contains, protention opens the very dimension of the future and thus makes full-fledged anticipation possible. In this sense, protention is the original awaiting of something to come. I will return on this topic in the fourth paragraph of this chapter. For the moment, it is important to underline the relevant role of protention in the creation of perceptual continuity.

* Differences between retention and protention.
Protention and retention differ in various respects. Firstly, they differ in the direction of their intentional movement and their intended object. Secondly, the protentional part of the temporal horizon is fundamentally different from retention in that it is largely indeterminate (because of its openness towards what is to come). Brough observes that “primary expectation thus differs from the other two intentional moments in the sense that its object, as future, is not yet determined. Still, as a mode of consciousness, it joins primary memory as ‘perceptual’ as opposed to ordinary or secondary expectation” (1991, p. XL). In Miller’s words, “what I retain at any given instant is retained by me as having (at least some) fully determinate extra-temporal characteristics; whereas, what I protend at that instant may suffer, and often does, from a uniform extra-temporal attributive indeterminacy. […] In a perceptual act, the meanings of my retentions are sensorily “filled”; whereas, the meanings of my protentions are “unfilled” (1984, p. 128-129).

This difference between protention and retention, as well as the asymmetry of retentioning and protentioning is well explained by Shaun Gallagher when he writes:

first, retentioning is always a fulfilled intentioning of previous functionings of consciousness which are actualized and determinate. In contrast, protentioning is an unfulfilled intentioning into an immediate but indeterminate about-to-become. The protended is pure possibility rather than something actual or actualized. There is nothing determinate for protentioning to protend since the intended about-to-be functioning of consciousness has not yet taken place. Thus, where retentioning is intentionally fulfilled, protentioning is intentionally empty. It follows, secondly, that whereas retentioning always involves a retentional continuum, protentioning cannot be a continuum, since the intended protentional functionings of consciousness have not yet occurred. At one point Husserl speaks of a “horizon of protentions” and this is preferable to what he refers to in Ideas I as a “protentional continuum” (1979, p. 455).

**How the three moments work together: structurality of the flux.**

We can now analyse how these three modes of the temporal structure work together in order to allow the perception of enduring objects with specific meanings (for example, the perception of a sequence of acoustic events as constituting a musical object). This quotation from Miller serves as a good point of departure: “according to Husserl, I have, at t, a primal-impression, a continuous manifold of retentions which are “modifications” of
past prima-impressions, and a continuous manifold of protentions which are responsible for my having, at t, an anticipatory awareness of a continuous manifold of future primal-impressions” (1984, p. 121). The basic idea is that we experience, in time (in the living present), a presentation not only of present events becoming past events, but also of future events becoming present.

Retention and protention interact in the articulation of the present, influencing its content. We do not consciously activate, nor are we aware of either of these modes, but by ‘working together’ they allow for the present perception of an object. Thanks to this interaction, a succession of sensory information is experienced as having continuity, and different events influence each others in a meaningful way. The experience is thus continuous throughout the flux of its impressions. The same idea of openness to the future previously discussed implies cohesion with past and present. In this way, the relationships between the three modes and their movement results in the unfolding of the horizon and its possibilities.

As we have seen, temporal movement is characterized by a synthetic hearing. This is a passive synthesis – in the sense that it is unconsciously realized – which allows the constitution of perceptual experiences in a process of continuous becoming. In the case of music, Smith notes that “in musical sound original impression, retention, and protention acquire a synthetic unity of their own in a temporal process of continual self completion independently of what we traditionally call conscious activity” (1979, p. 108). This movement allows listeners to always have new and subjective meaningful experiences while hearing music: “as past, present and future overlap and conjoin, their interaction in one’s consciousness opens endless possibilities for musical meaning. Each listening can be new and fresh and still be grounded in previous listenings. The past continues to be adjusted in light of a future that cannot be precisely predetermined” (Ferrara, 1991, p. 149). Through the living body and its interaction with the environment, sound events are related into a unitary whole and the just-having-been and the just-about-to be are tacitly retained and protended in order to constitute meaningful experiences.
Temporal flux is thus characterized by a structurality which is developed in the living present. As we have seen, the ‘now’ is intrinsically spanned within itself thanks to the inter-play of retentions, primal impressions and protentions. The succession of events is organized along a continuum: Louis Sandowsky writes that “retention gives the passing over of the present into that which is no longer (pastness), while protention fills the present with the anticipation of that which is not yet (the pre-expectational horizon of open futurity)” (2006, p. 63). The mind not only retains the impressions that pass-away, but is also already at work anticipating their arrival: this shows a structural way of working which is fundamental for the consciousness of successivity.

This point supports the ideas discussed in the previous chapter, namely that the human mind works structurally in order to have musical experiences. In particular, through dynamic temporal movement listeners structurally protend and retain information: this is fundamental for the creation of relationships between sounds and their perception as a unitary whole. As Gallagher observes, “Husserl describes direct experience as something organized and structured, i.e., structured by the functionings of retention, primal-
impression, and protention” (1979, p. 448). Every experience is thus phenomenologically structured in the impressional, retentional, protentional pattern. This is the basis for every perceptual act.

At a more complex level of the temporal structure – which will be discussed forthcoming – listeners relate sounds into musical structures: this allows the constitution of musical forms during the perception of sound events occurring in the auditory environment. The temporal structure represents an unconscious and universal ability through which musical experience is organized and structured. In this case, expectation and memory are involved as fundamental processes of the perceptual act that accounts for the awareness of musical meanings. An example that well describes this structural and temporal movement concerns the perception of a tonal cadence or the grouping of three chords into a meaningful unit.

\[
t1 \rightarrow C \ I \ retention \\
t2 \rightarrow G \ V \ primal \ impression \\
t3 \rightarrow C \ I \ protention
\]

\[\text{Figure 10a.} \text{ Tonal cadence in C major.}\]

\[\text{Figure 10b.} \text{ Grouping of three chords into a tonal cadence.}\]
An experienced Western tonal listener perceives a cadence because she relates and groups sounds in a particular way. In order to be aware that she is perceiving a cadence (that is, an auditory event that acquires the meaning of ‘cadence’ for the subject who experiences it) and in order to structurally organize sounds in the form of a cadence, she firstly has to relate the acoustic events that are present to her consciousness (during the living present) into a unitary enduring object. Without this ability she would not have a basis for forming relationships and constituting meanings, and there would not be any continuity between the events occurring in time (they would be isolated events without any kind of temporal direction). In other words, there would be no directionality of the temporal flux. During the present moment – the now-phase of the musical experience – the listener is able to relate the three notes or chords (C, G, C) into a unitary whole thanks to the temporal movement of retention, primal-impression and protention. She experiences the C, G, C as
belonging to the same object, which will acquire a specific musical meaning (that of a cadence or a musical unit) because of a structural organization that is realized at a more complex level of the temporal structure and as a result of certain perceptual processes. At this level the processed involved allow for the segmentation and grouping of sounds into chords that are structurally related in the form of a tonal cadence ending a musical passage. In this way, a meaningful musical experience is constituted because the sounds occurring in the acoustical environment are perceived as having a musical meaning.

Returning to what I consider to be the basic level of the temporal structure and making reference to the example in Figure 10c, we can see that at $t_2$ the listener experiences G under the temporal mode of primal impression, but she also retains the C she experienced at $t_1$ as just-having-been, or as having elapsed in the past but still present in the now-phase thanks to retention. This is subject to modification because of the new phase, the G of the primal impression, but, at the same time, it influences the experience of the G in the now-phase. Moreover, at $t_2$ the listener also has a protentional experience toward a future tone-phase, and she has an anticipatory awareness of a continuous manifold of future primal-impressions. In the characteristic openness of protention and the indefinite possibilities offered by the horizon, at $t_2$ she protends toward $t_3$, which could be C in the case of a tonal cadence. This does not mean that she necessarily protends toward C, but this is one of the possibilities she has. Primary impression is also the fulfilment of what was a protention: at $t_3$ the protention the listener had at $t_2$ is fulfilled. This example involves both the temporal structure of consciousness (its basic and complex levels) and the structure of a musical object.

The listener’s perceived structure is developed through the structure of the temporal movement which characterizes human consciousness and awareness of one’s own perception. The basic level of the temporal structure (the movement of retention, primal-impression and pretension) is the fundamental experiential ground on which more complex organizational and perceptual processes intervene in order to constitute meaningful forms and allow the listener to be aware of her auditory perceptions and the objects she experiences.
Some critiques of Husserl and his theory.

With the help of Gallagher’s work *Suggestions Towards a Revision of Husserl’s Phenomenology of Time-Consciousness* (1979), in this paragraph I will highlight some points related to Husserl’s account of time which need clarification. The first aspect of Husserl’s account of the temporal structure which is not clear is the distinction between protention and expectation. While he wrote more about the role of retention and its differences with secondary memory, he did not develop an analysis of expectation, despite some references to it. In particular, it is not always clear whether Husserl feels that protention and expectation are both acts of consciousness. Gallagher makes an interesting suggestion: the elements of the temporal structure could be understood as functions that allow the constitution of musical experiences and the creation of structural relationships between sounds in order to perceive music. More precisely, he critiques the concept of phase and Husserl’s idea that for each note of a melody there must be a perceptual phase that includes a primal-impression, a retention and a protention in order to intuit the temporal position of the present-note in relation to the other (past and future) notes of the melody. Retention, impression and protention are not individual phases of perception because a phase is not an enduring ‘extension’, while perception is precisely an enduring ‘extension’. In Gallagher’s view, primal-impression, retention, and protention are not temporal phases of the flux, but structural functions that allow the organization and perception of the continuity of a same object.

Gallagher also underlines the need to specify Husserl’s concepts: “the concepts of prima-impression, retention, and protention are most fundamental to Husserl’s phenomenology. Nonetheless, they are not the clearest or most precise. Retention is called ‘retention’ as many times as it is not – i.e., it is referred to as ‘consciousness’, ‘phase’, ‘memory’, etc. Primal impression is termed ‘consciousness’, ‘sensation’, ‘perception’, etc. And protention, a concept not completely developed by Husserl, is termed ‘intention’, ‘fore-seeing’, and ‘expectation’” (1979, p. 452). Husserl is ambiguous about the way these moments have to be considered: are they acts of consciousness? “Is retentional
consciousness a consciousness in the same sense as perceptual consciousness is?” (Gallagher, 1979, p. 452). For example, he is vague in saying that retention is a mode of consciousness. “Retention, since it is not an act, has no act-character and so it cannot be a mode of consciousness in the usual sense” (Gallagher, 1979, p. 453). If retention is not an act or mode of consciousness, it may be understood as a performance, which essentially is a function of consciousness. In particular, the function of retention is to retain the just-having-been consciousness.

Protention is more problematic. Gallagher points out that “Husserl fails to develop a description of protentional functioning beyond a few brief indications. The most significant of these is that protentioning is a prima-expectation. Similar to retentional functioning, protentioning is not a conscious act of expectation – it does not thematize or make an object of what it intends. It is a tacit intentional performance of consciousness whereby the immediately intended horizon of consciousness yet to come is integrated with the speciously present consciousness” (1979, p. 455).

Authors have also highlighted that in his account of time-consciousness and its temporal structure, Husserl did not develop an in-depth description of protention, which is mentioned infrequently in his work. As Lohmar observes, if we compare this with his analyses of retention, protention seems to be mentioned only for reasons of symmetry and fairness, despite retention and protention being distinctly dominant both quantitatively and qualitatively.

As previously mentioned, I argue for the idea that protention is not an act of consciousness, nor is it a conscious act that makes the listener aware of her experience of an object. It follows that it is not a perceptual act, but it forms (together with retention and primal impression) the grounds for perceptual awareness of the objects that are present to consciousness. It is a constituent part of time-consciousness (previously defined as the primordial place for the constitution of objects and the origin of meaningful forms) because it is an essential part of its structure. It contributes to the creation of relationships between

---

118 For an overview to the problem see Lohmar (2002); Rodemeyer (2006).
events that are at the basis of the perception of enduring objects, makes events present to
the consciousness and contributes to their integration into unitary wholes.

As part of the basic level of the temporal structure, protention allows the experience
and subsequent perception of objects, and it is itself experienced in consciousness but not
perceived. This means that even if they are functions of consciousness, we are not aware of
our protentional movements, while we may be aware of our anticipations and expectational
states. We are not aware of our functions and structures, as they are not conscious acts, but
constituents of the temporal flux of consciousness.

Gallagher recognizes that “the concept of protentional functioning obviously
requires further clarification, and an interrogation of experience might well be continued
with a view to providing it” (1979, p. 457). My question is: could analysis of human
perception within the field of cognitive psychology be helpful in order to provide an
explanation of this concept? A related question which exemplifies the problematic gap
between empirical science and phenomenology is: could objective empirical studies be
helpful in explaining the structure of an essentially subjective experience such as musical
understanding and the constitution of musical experiences?

**Conceptual distinctions: a renewal**

Based on Husserl’s account of the temporal structure, I propose to further develop
the distinction between expectation and anticipation which I introduced in the fourth
chapter. More precisely, I suggest a clarification of the role of protention, expectation and
anticipation in the constitution of subjective musical experiences. Such an attempt may
provide interesting insights for the comprehension of the man’s musical experience in the
metro – particularly in the case where he experiences unfamiliar music.

The temporal structure of human consciousness seems to develop itself at two
levels of complexity. At the first level, due to the movement of retention-primal
impression-protention, the listener is able to perceive a sequence of acoustic events as an
enduring object. Such events are thus experienced as a unity persisting in time. At the second level, memories of past experiences and expectations about the future are involved in order to constitute the meaning which that experienced object has for the subject. Both levels of complexity are part of the perceptual act and, thanks to the contribution of other cognitive principles, they allow for a meaningful organization of sounds in time. These two levels are strictly related: secondary memory and expectation elaborate and solidify what the more primitive forms first make available. In this way they realize a basic perceptual understanding (which is a musical one in the case of the man in the metro).

Bearing this two-level structure in mind, the questions that I address in this paragraph concern the main differences between protention and expectation: for example, what are protention, expectation and anticipation respectively directed towards (is it something specific, or an indefinite future event which is contained in the possibilities offered by the horizon)? Are the three phases forms of temporal awareness or appearance? What kind of mental processes do they entail?

Expectation and the attempt to anticipate the future are explained through the concept of the horizon, which is part of the temporal structure and represents the place where intentional movement toward an unknown future is developed. At the basis of this concept is the idea that, as Miller writes, “objects are perceptually experienced by us inadequately”: at any given moment the object is always experienced by us as from a certain perspective, as having “more” to it than is captured by our perceptual act at that moment” (1984, p. 82). We have an incomplete perceptual experience which involves an intentional movement toward the possibilities that could fulfil the content of our perceptions.

Husserl calls the features associated with the act which ‘complete’ and prescribe the course of experience ‘the horizon’ of the act. The horizon is thus understood as a set of possibilities to which the consciousness points during the perceptual temporal experience. It guarantees the subjective aspect of the experience, which becomes clear in the process of constitution of meaning. It also shows the influence of learning and past experiences.¹¹⁹

¹¹⁹ As Miller observes, one’s past experience plays an important role in determining the horizons of their perceptual acts. “Another part of the horizon of my act is constituted by my memories from my earlier
Miller is clear in underlining this aspect:

…these “predelineated potentialities” of further experiences which would “complete” the present one are not (necessarily) those possibilities which exist “there”, “objectively”, as a mere matter of metaphysical fact. They are, rather, possibilities of further experience which are in fact “projected” by the agent, possibilities, which may or may not become actualities for that agent as the experience progresses. These “predelineated potentialities” are experienced by us (“projected” by us) through expectations and memories associated at any given moment with our perceptual act. […] Indeed, to say that we experience the object as having “more” to it than is captured by our present perception is, in part, to say that we expect other features of that object to reveal themselves, given a suitable observational standpoint. How rich those expectations are depends, in part, on our past experience of that particular object, or of objects of its kind, as well as on our general background beliefs (1984, p. 83).

The horizon of the act of perception, then, consists of a pattern of recollections and expectations regarding past and future experiences in relation to the present act of perception, whose content continually changes as the perceptual experience progresses. Retentions and protentions are associated, at any give moment, with the act of perception. They represent two other parts of the horizon, but they do not coincide with expectation and memory because their function is different. Miller goes on to note that “[r]etention and protention are technical terms in Husserl’s theory of temporal awareness, and they are used by him to designate those features of the act which at any given moment are responsible for our awareness of the temporally adjacent parts of our experience – those occupying the immediate past and the immediate future relative to that moment” (1984, p. 84).

Memory and expectation are part of the general horizon of the perceptual act, and may be considered as acts of consciousness that refer to specific contents. For example, the content of an expectation is a future possibility about how the perceived temporal object might develop. Retentions and protentions, rather, are part of the horizon of the act in which an object is constituted in consciousness as a unity. Protention refers to an indefinite future, but it is fundamental in creating the essential link with what has to come: without it, connections between one perceived event and the next are not possible.

---

120 experiences (if there were any) of the same object. These recollections constitute the past-oriented part of the horizon of my present act” (1984,p.83).

120 As Miller writes, “[e]xpectations belonging to the horizon may be characterized either as expectations regarding future perceptions of the (currently perceived) object, or as expectations regarding features of the object itself other than those currently being presented” (1984, p. 86).
While retention and protention allow for the continuity and sameness of present perception, memory and expectation enable the constitution of the subjective meaning which that perception acquires. The following quotation from Moran’s *Introduction to Phenomenology* (2000) well explains the concept of horizon and shows its complexity:

Husserl recognises not only that in any perceptual act there is present the actual side of the object perceived, but also that each act of perception takes place within a horizon of anticipations. I know I will be able to see other sides of a table if I walk around it; it will be resistant to touch, I may not be able to lift it, and so on. Subsequent perceptions either confirm these anticipations or else set up a whole new domain of anticipations. In our ordinary experience we are interested in the confirmations themselves; in phenomenology we are attempting to delimit the nexus of expectations etc. which are enabled by the perceptual act as such. Husserl is particularly aware that perception is a temporal process; it does not take place wholly in the present but is oriented towards future experiences and at the same time is an experience of enduring or continuing from past experiences. There is also a “horizon of the past”, the potential to awake recollections. Crucially, for phenomenology, Husserl was alert to the fact that many of these unrealised possibilities, which are given ‘horizontally’ in any experience, are given in the form of possibilities which I myself can carry out. They have the character of “I can”, as Husserl says (2000, p. 162).

Based on the concept of ‘horizon’, we can examine the first important difference between protention and expectation. As we have already seen, protention, for Husserl, is the openness of the consciousness to the future. It is indefinitely directed toward the future and the only definite thing is that without exception, something will come. In other words, it is an indefinite openness to the indefinite possibilities of the horizon. Protention is thus anticipatory in the broadest sense of being directed towards the just-about-to-be, but it does not have the structure of the act of expectation. It protends to a pure possibility, immediate but indeterminate, not yet taken place, empty. Protention is only determined by what one is experiencing in the now. Brough writes: “If I am presently hearing a symphony, my retentional consciousness will intend elapsed notes of the symphony as they elapse, and my protentional consciousness will intend phases of the symphony yet to come. As long as I continue to hear the symphony, these protentions will be fulfilled” (1989,p. 278).

Expectation and anticipation are part of a general tendency of human beings to plan the future. In generating expectations and anticipations, individuals expect their plans to come to fruition. In particular, expectation refers to a range of probabilities within the indefinite possibilities of the horizon. As we have seen through analysis of Huron’s theory
about the expectation process, such probabilities are built upon a hierarchy of possibilities that the subject maintains on the basis of a statistical learning during her exposure to the environment.

Within Huron’s framework, I suggested that anticipation refers to a very specific event: the most probable occurrence within the possibilities of the horizon and the probabilities offered by immediately previous expectations. It follows that, in relation to the degree of certainty about the future, the protentional movement does not involve any degree or probability because it is a tendency toward the future that simply allows for the establishment of connections between successive events in time. It is not directed toward some probable events, but to all the possibilities contained in the horizon that may fulfil the content of our perceptions. Its essential function is to contribute (together with retention and primal impression) to the creation of those relationships that form the basis for the perception of enduring objects.

In the case of expectation there is a more or less precise degree of certainty, which depends on an individual’s knowledge and past experiences with particular events. If an expectation is very strong, it leads to the anticipation of a specific event. In this case, the degree of certainty is very high and the subject knows what will happen because, for example, she had previously experienced it.

The probabilistic character of the expectation process that has been presented by Huron is thus particularly relevant in order to make a distinction between protention, expectation and anticipation. While protention is a function of the temporal structure that allows the relation between successive moments and events in time, expectation refers to those events which are considered as the most likely to occur in the future, and anticipation indicates a sort of anticipatory perception of the event one expects to occur. The probabilistic and statistical character of the expectation process – which in Huron’s writing acquires a biological meaning – becomes clear through the phenomenological concept of the horizon. Thanks to the second level of the temporal structure, listeners recall schema that are based on past experiences and, on that basis and in relation to the present, they create expectations about future events or anticipate specific events. Again, the role of
memory is fundamental because it informs future events, while relation to the present is essential for the creation of expectations and the belief that what is expected will exist: expectation is consciousness of what is absent only through something present and the information received from memory of past experiences.

Listeners, during their experience in the world, develop a statistical hierarchy of events that guides the pattern of probabilities. This helps them, for example, to predict the likely temporal placement of future tone onsets. This movement has a subjective character because the creation of a hierarchy of probable events is based on an individual subject’s knowledge and experiences. In the case of very familiar music, listeners can have quite precise anticipations due to the contribution of learning and memory. All this also indicates that through expectation – the going forward to intend the now that is not yet and which grows out of our past experience – the subject, in her experiencing the world, anticipates the future by a protentional having-in-advance of future contents. In this way, perception also becomes an interpretation of the experience and the objects intentionally ‘merge’ in their specific features and contents. The constitution of meaningful experiences is thus realized.

Considering the object of reference, it is clear that in the case of music, protention is directed toward the next sound in a succession of musical sounds (being it a melody, an atonal composition or a serial piece of music) which is not defined, since it belongs to the indefinite possibilities of the horizon. As part of the triple intentionality of the temporal structure, protention has the goal of establishing continuity between different sound events. It is responsible for the experience of events as succeeding, connected and temporally related into a unity. There is no awareness of any intentional movement toward the next sound. Protention is an experiential movement that does not involve the awareness of meaning, but it is the necessary basis for the constitution of meaning. It is a function of consciousness that relates events in time.

An objection could be made here. I defined protention as an experiential movement, or as part of the experience of the objects that are perceived by human consciousness. In the first and second chapters I underlined the phenomenological idea that ‘the experience is
already meaningful’. Through experience and thanks to the temporal structure, succeeding events are related and presented to the consciousness as unitary objects. As objects of consciousness they already have a meaning, of which the subject becomes aware thanks to the perceptual processes that are involved during the perceptual act. Protention contributes to the experience of objects, but is itself not experienced. The subject cannot be aware of protention, nor perceive it, and it does not carry any specific meaning. It is not an object or an event, but a structure of consciousness having a specific function.

While protention is characterized by the awaiting of indeterminate events (for example, an acoustic event in the case of music), expectation and anticipation involve awaiting a range of probable events or specific events that relate to both present and retained experiences (for example, the note C at the end of a cadence). The expected or anticipated events may be considered as objects of consciousness because they have a meaning in themselves (that of being a cadence, the last movement in a sonata form, a C chord). However, they are not present to the consciousness yet and they are not sensorily experienced. For this reason, these expected or anticipated events are not perceived but they influence the meaning of the perceived object – which, as we have seen, cannot be perceived in a single moment.

Expectation and anticipation, whose intentionality is directed toward future events in order to prepare the organism to respond appropriately or to act before these future events happen, respectively contribute (together with other cognitive processes) to the creation of structural and meaningful relationships between sounds. Both influence the organization of experienced sounds into structural forms, which acquire a particular meaning for the subject (such as that of being a ‘cadence’). For example, on the basis of segmentation and grouping and creating patterns of tension and relaxation, expectation and anticipation determine when a musical sequence or movement finishes and another begins. In this way they allow the constitution of meaning because they connect the events that form the experienced object in meaningful ways. The expected or anticipated events become more complex than in the case of protention: they are not only the immediate events to come, but meaningful events. For example, after the first movement in C major I
anticipate that the attack of the new movement will be in G major, or I simply expect that the movement will continue in the same key or with a variation in timbre. In this way expectation and anticipation constitute events that belong to the same object in a meaningful structure. In this sense they are part of the perceptual act.

To conclude, protention is a standing condition for experiencing music, but it does not determine what is specifically heard and the meaning it acquires (music instead of the noise of the metro, a pop song instead of a jazz performance). The question about how we judge the end of one sequence and the beginning of another (and thus structurally organize experienced events and perceive them as meaningful objects) has to do with our expectations and anticipations, as both influence our awareness of the meanings we constitute during the perceptual act.

Thus, while retention and protention allow for the continuity and sameness of present perception, memory and expectation determine the constitution of the subjective meaning that that perception acquires. The next problem involves better understanding of what anticipations are and in what ways they differ from expectation. In order to accomplish this, I take into account the fundamental question that I introduced in the fourth chapter: what kind of mental states are protention, expectation and anticipation?

An important difference between protention and expectation is that the latter is considered, by phenomenology, as an act of consciousness rather than an experienced mode of temporal appearance, even if it is part of the temporal structure at a higher level of complexity. Brough’s claim makes this point clear: “if perception is the consciousness of what now exists as present in person, memory is the consciousness of what is past and expectation is the consciousness of what is future” (2005, p. XXXIV). While protention is a form of temporal appearance which is part of the experience of an object in that it allows perceptual awareness of it, expectation and anticipation are both acts of consciousness. They differ first of all in their object of reference: while in expectation it is not well defined and belongs to the horizon of future probabilities, in anticipation it is something specific and already known. Expectation may be understood as an action tendency toward the future, as representing a dynamic movement which is open to the changes occurring in the
environment during the temporal unfolding of experienced objects. It belongs to the field of intentionality and it indicates that the perceiver is not a passive receiver of uninterpreted data.

In his study of Husserl’s phenomenology of time-consciousness, Brough is clear in making a distinction between protention and expectation that explains the different structurality of protention as compared to the structure of the act of expectation:

What is intended in protention appears as not yet now, as future. Protention is not expectation, which is a fully developed representational act, any more than retention is secondary memory. In expectation, a future object or event is run through in substantial part or even from beginning to end as if it were present. In expectation, the object will appear with represented now, past, and future phases, and therefore will involve represented primal impression, retention, and protention. In the protention which belongs to the actual momentary phase of consciousness, however, only what is future is intended, and it is intended or given precisely and only as future, as moving towards the now (1989, p. 278).

If we consider expectation – as opposed to protention – to be an act of consciousness and not a mode of appearance, we might define it as a pre-presentation of a future event but not a present perception of it in the same way in which Husserl defines memory as a re-presentation of an object that has been perceived in the past but that does not coincide with a new perception of such object.

Expectation is a cognitive process and a temporal movement that develops itself during the perceptual act, but it is not in itself a perception because the event is not present nor is it clearly defined. In Brough’s opinion expectation is a fully developed representational act. However, empirical studies in music show that expectation refers to a state of tension about what has to come, but it is not a conscious representation of a specific expected event. As we have seen, in expecting a future event there also is some kind of reference to its being. This could delineate a difference between expectation and anticipation. Expectation refers to something possible and belonging to a range of probabilities within the indefinite possibilities of the horizon, while anticipation attends to a specific event in the future.

If protention contributes to the presentation of an experienced object and expectation is a pre-presentation of a possible, not clearly defined future event belonging to
a range of possibilities contained within the horizon, could anticipation be considered as a
‘quasi-perception’ of a future event that we ‘perceive’ in advance and without its sensory
occurrences in the moment of its anticipation in our mind?

I suggest considering anticipation as a mental representation that may lead to the
awareness of a mental projection about a future predicted event. It is the action of mentally
representing a future expected event or a known outcome before its occurrence in time and
after a strong expectation. It is not an act of perception because it does not coincide with the
sensory experience of the anticipated object (which is not present to consciousness) and it
does not have the complex structure of a perceptual act. The man in the metro does not
have any sensory sound experience of what he anticipates in the ‘now’.

I propose to consider anticipation as a ‘quasi-perception’ because it may coincide
with the awareness and a kind of mental representation of the specific object or event that is
expected. The man in the metro, for example, is aware of the immediate occurrence of a
musical event (for example, the introduction of a busker’s violin playing a particular
melodic line at a specific moment in the performance) as if that event is out there, but he
does not perceive it in that anticipatory moment. Such anticipatory representation is built on
the basis of past experiences, the knowledge of the event, the context in which it occurs and
the familiarity that the perceiver or listener has with it.

In the case of expectation, on the other hand, the perceiver can be conscious or
unconscious of her expectations. The vast majority of expectations are unconscious; they
are occurring all of the time. Reading a sentence provides a good example, since the mind
is forming expectations throughout the reading process, but it is rare that our sense of what
words should occur next rises up into consciousness. Notice that I am not talking about the
specific words as they are read, since clearly the words do appear in consciousness. I am
speaking of our awareness that we form expectations related to the sequence of words. We
are rarely aware of such expectations while reading. A perceiver can, however, be aware of
experiencing a state of expectation and suspense about something in the future which is not
exactly known. Because of such uncertainty, the state of expectation does not correspond to
a specific mental representation of which she can be aware. It is a kind of pre-presentation
of future possible events which differs from a re-presentation in that there is no certainty about the real occurrence of the expected objects or events (while in the case of memory the object is already known and it is re-presented in all its aspects). In this sense expectation is more related to a mental state, while anticipation is a kind of action that corresponds to a specific mental representation. A problem with this account concerns the way mental representations have to be considered.

In general, a mental representation is an image, a concept or a belief. It has a strong subjective character which depends upon one’s own experiences. In the case of anticipation in music, the related mental representation (conscious or unconscious) has as its content a highly expected future event which the listener has in some way experienced in the past. It could correspond to a structural feature of the music, a particular instrument, a meaning or musical content which is related and determined by the subject’s experience. Thus, we have mental representations of music because we refer it (in the form of an auditory image, content, concept or belief) to something that belongs to our experience. We relate a musical event to something which is or has been part of our subjective experience, and anticipation has to be understood in this framework. This underlines the subjective and experiential character of the mental representation that a listener forms in relation to a musical object and in order to anticipate its future ongoing.

I do not discuss the problem of the specificity of expectation – what a listener expects at any moment of the musical listening – because this is a task for the empirical sciences. In music, for example, a listener may expect the pitch C4 to occur on the next beat. However, one might argue that this description of the anticipated C4 is not ‘specific’ enough. By ‘C4’ the listener might expect the pitch C4 plus/minus 25 cents; and by ‘next beat’ the listener might expect the tone onset to be within a window of 125 milliseconds. Wouldn't we also have to say that the listener expects a ‘trumpet’ or a ‘piano’ playing C4 on the next beat? And even if the listener expects a trumpet, there are many different trumpet timbres. In other words, when we say ‘C4 on the next beat’, one researcher might claim that that is a ‘specific expectation’ whereas another researcher might claim that that is
a ‘whole class of possible events’. In this way, it becomes difficult to say whether the listener is ‘anticipating’ C4 on the next beat or merely ‘expecting’ C4 on the next beat.

In my opinion, this is due to the subjective character of the musical experience and the influence of the context and situation. The specificity of anticipation is a subjective matter: it is something that cannot be entirely predicted or explained through an empirical investigation. Because of the subjective character of the musical experience it is not possible to exactly define expectation and anticipation (and their intentional objects). The specificity of a subject’s expectation or anticipation may refer to different objects, contents and mental representations. What I would like to underline is that expectation and anticipation influence, in different ways, a subject’s musical experience and her responses to a sound event (which, in both cases, can be positive responses). This might be studied by observing the listener’s bodily responses and judgements of unexpected events such as changes in timbre, dynamics or rhythm.

Expectation and anticipation are both part of the human intentional movement toward the future, but they may be understood as two slightly different cognitive mechanisms. Even if they are strictly related, they may use different cognitive structures during perception. It would be interesting to study such a possibility by analysing the different bodily and emotional responses individuals have when they listen to more or less familiar music. I hypothesize that such responses (action-readiness, in Huron’s words) are different in cases where listeners anticipate or simply expect future sound events. For example, expectation prepares the subject to respond to some probable events, while anticipation allows the subject to act in response to the anticipated event before it occurs.

It follows that expectation and anticipation influence the constitution of musical experiences and the subjective meaning one’s sound experience acquires in different ways. In the case the man from our example is very familiar with the music he hears in the metro, he is able to anticipate. His experience acquires a particular meaning in relation to his past experience of that same music and the context in which he actually hears it.

The occurrence of unexpected or non-anticipated events can result in different responses (for example, different emotional experiences). In the case where the music is not
familiar, the man in the metro tries to create expectations on the basis of his knowledge and the sound relationships he creates during the present perceptual act, but he is not able to anticipate the ongoing of the musical event. This might result in a sense of surprise when his expectations are not satisfied and/or it may result in an interest for the new music. He might also have a sense of misunderstanding if the music is totally unfamiliar or the event is completely unexpected. In this case, the lack of anticipation is one reason (but not the only one) for the sense of misunderstanding that people have when they listen to unfamiliar music.

**Listening in the moment.**

I began this study by addressing the question of what it means to perceptually understand the auditory environment as music (i.e., being aware of musical events occurring in the acoustic environment), how sounds are constituted as music by the listener and what role the process of expectation plays in the constitution of musical experiences. I propose at this point to relate the phenomenological account of temporal structure to what I have called (under the influence of Levinson’s book *Music in the Moment*) ‘listening in the moment’, and show its essential contribution to basic musical understanding. To hear music or, better, to perceive music within the auditory environment (often very complex, as in the case of the metro), listeners do not necessarily need specific musical knowledge or the ability to perceive music in a tonal manner. They basically only require the ability to create relationships in time, to organize sounds on the basis of their learned schemas and to relate some acoustical features of sounds. In this way, listeners realize the intentional movement of human consciousness.

As we have seen in the previous section, expectation and anticipation play a crucial role in determining the meaning that one’s individual experience acquires. I now argue for the idea that the temporal structure makes possible a kind of comprehension of unfamiliar music and it represents the way through which human beings become familiar with the
living world. I discuss this ability of the human consciousness taking into account Levinson’s Concatenationism view and relating it to the first level of the temporal structure previously analysed (the movement of retention-primal and impression-protention).

As noted in the first chapter, concatenationism underlines the moment-to-moment relationship between musical elements during the unfolding of a musical structure. Music perception is seen as a chain of interconnected and interpenetrating links. Discussing how individuals listen to music, Levinson proposes the phenomenon of following music as fundamental in order to have musical experiences. This describes an involvement in the movement and progression of music moment by moment. In this context, music is understood as a chain of overlapping and mutually involved parts of a small extent, rather than either a seamless totality or an architectural arrangement. In Levinson’s words, “musical understanding centrally involves neither aural grasp of a large span of music as a whole, nor intellectual grasp of large-scale connections between parts; understanding music is centrally a matter of apprehending individual bits of music and immediate progressions from bit to bit” (1999, p. 464).

The issue at the basis of Levinson’s work is that “music of any extent consists of a series of successive events, which cannot be apprehended simultaneously in a single perceptual act” (1997, p. 2): what is crucial for the comprehension of sounds as music is the local movement from note to note and phrase to phrase, and the involvement in the musical progression from point to point. From this point of view, the experience of music is fundamentally a matter of individual momentary impressions that listeners are able to relate into unitary wholes thanks to the basic level of the temporal structure, which is at the basis of every perceptual act.

With reference to the perception of a melody, Husserl supports the idea that listeners can hear only the actual sensorily experienced tones of the melody, and not the melody as an object at any given instant:

Let us take a particular melody or cohesive part of a melody as an example. The matter seems very simple at first; we hear a melody; i.e., we perceive it, for hearing is indeed perception. While the first tone is sounding, the second comes, then the third, and so on. Music we not say that when the second tone sounds I hear it, but I no longer hear the first, and so on? In truth, therefore, I do not hear the
melody but only the particular tone which is actually present. [...] Every tone itself has a temporal extension: with its continued sounding, however, it has an ever new now, and the tone actually preceding is changing into something past. Therefore, I hear at any instant only the actual phase of the tone (1964, p. 43).

The experience of continuity during the perception of a musical event is guaranteed by the basic level of the temporal structure. As we have seen, the temporal movement of retention, primal impression and protention allows the perception of successive events as belonging to the same temporally extended object. Temporal movement is thus at the basis of one’s awareness of musical objects, which could not otherwise be apprehended in a single momentary phase.

In this way, Husserl seems to present the condition for every kind of musical understanding and constitution of musical experiences: without the ability to relate succeeding sounds into unitary wholes one would only be capable of experiencing acoustic events as single and disconnected sounds. In Levinson’s words, “hearing musical movement is necessarily hearing a sonic entity not all of which is sounding at any instant, while at any instant, one hears the sounding notes as belonging to a musical flow, or as contained within a musical process, of which they form a part” (1997, p. 15). This quotation clearly shows that notes (sounds) are related during the flow of the present perception. It recalls the way the three modes of temporal appearance (retention, primal-impression, protention,) work together: the listener is focused on the connections between individual occurring parts with immediately preceding and succeeding parts. In this way, she perceives music as a developing process which is characterized by continuous changes and modifications. As we have previously seen, each moment influences and is modified by the others.

The ‘basic’ temporal structure thus allows the development of perceptual acts and the awareness of what one is perceiving. It is fundamental to basic musical understanding and the perceptual awareness of musical events occurring in the auditory environment. Levinson relates the concept of ‘quasi-hearing’121 – which is central to his theory – to the

---

121 ‘Quasi-hearing’ is a special kind of perception, defined as the “experiencing of a stretch of music as if one were hearing it in its entirety, while actually only hearing some bit contained within that stretch. It involves apprehending a stretch of music as a motion and yet complete, or as an almost-present whole, through the
temporal structure, making a clear parallel with Husserl’s phenomenology of internal time-consciousness through the following reference to Clifton’s work *Music as Heard* (1983):

The experience of quasi-hearing can be usefully thought of as having three components or aspects. The first would be the actual hearing of an instant of music, the second would be the vivid remembering of a stretch of music just heard, and the third would be the vivid anticipation of a stretch to come. Vivid memory and vivid anticipation might be thought to provide tonal images that exist for listening consciousness simultaneously, yet somehow like the way the peripheral objects of vision are present to the eye, though obliquely, at the same time as the object that is in central focus. The width of the window of quasi-hearing, so conceived, is thus at any point a direct function of the reach of vivid memory and vivid anticipation at that point, which is a matter of the extent of virtual imagining backwards and forwards that the musical material and one’s familiarity with it allow (Levinson, 1997, p. 16).

Retention corresponds to ‘vivid memory’ and protention to ‘vivid anticipation. The ‘now’, or the present listening, is a quasi-hearable span surrounding a given instant. The listening horizon is an extent of music beyond what is quasi-heard and provides a sort of context to that which is. An interesting aspect of Levinson’s quotation is that anticipation seems more related to an auditory representation – in this case, a tonal image – than a protention in the phenomenological sense. This stresses the involvement of the second level of the temporal structure, which is fundamental to the constitution of meaning in music.

At first sight, Husserl’s description of the way successive events are apprehended as part of the same object (i.e. through the movement of retention, primal impression and protention) seems to correspond to Levinson’s description of the phenomenon of following music and its essential role in the perceptual comprehension of music. Levinson observes that “the momentary parts of music regularly group themselves into extended units that we seem to grasp as a whole, though we cannot literally perceive them at one time” (1997, p. 4). The reference to this grouping process makes an important distinction between the two philosophers and allows us to define the function of the two levels of the temporal structure.

Music is a process in time, characterized by a succession of impressions that are related into a same and continuous (durational) unit. Retention and protention allow for the development of music and the relation of parts into a meaningful whole. The musical form offices of vivid memory and anticipation” (Levinson, 1999, p. 474).
itself is centrally a matter of cogency of succession, moment to moment and part to part relationships, but the structure it acquires depends on organizational processes that are developed within a temporal movement that involves memory and expectation.

While retention, primal impression and protention represent the conditions for relating sounds in time, memory and expectation determine how we relate them in order to constitute meanings. For example, they allow us to understand why something ends in one way or the other, or why some patterns are used instead of others. Expectation, anticipation and memory operate on the basis of some cognitive organizational principles (such as similarity, grouping, and segmentation) which contribute in a strong way to the constitution of meaning. Because they themselves are cognitive processes, studies in the field of music cognition and perception help us to better comprehend the way sounds are related in time and why they are related in one way or the other.

Returning to the comparison between Husserl and Levinson, we can conclude that the basic level of the temporal structure is fundamental in providing a sense of unity and coherence during the unfolding of a musical piece:

The experience of unity or coherence I am invoking has as its basic content the sense that a passage currently being heard fits with and shares something with those that have preceded it, that in coming where and when it does, a given passage in a strong sense continues, as opposed to just succeeds, what has transpired up to that point. So construed, it is clear that the crucial experience of unity in a piece of music is one that can be had as the piece proceeds, without either intellectual grasp of widely separated parts or explicit conception of the whole (Levinson, 1997, p. 59).

This quotation presents the problem [not considered by Husserl] of the specific modes or processes which allow for a unified experience of music. More precisely, in a complex perceptual environment like that of the metro station, what is it that determines the experience of unity between successive events? What is it that determines the connections made between some events instead of others that, because of such a connection, become part of a unitary whole? I assert that the man in the metro relates certain (musical) sounds together – thus separating them from other acoustic events – because of their physical properties; because they share the same sound source and spatial location. Cognitive principles such as similarity and sound localization are also involved at this basic level of
the perceptual experience. Because the acoustic events that have been distinguished from
other events on the basis of their properties become part of the object which is present to
the consciousness as having a particular meaning (that of being, first of all, music), the
necessary ability involves relating them in time.

The temporal movement of retention, primal impression and protention that is at the
basis of ‘listening in the moment’ is what enables one to relate acoustic events in time.
Together with the second level of the temporal structure, which is based on past
experiences and knowledge, such temporal movement allows the man in the metro to
perceive music and to constitute his awareness of a temporally extended object as having
the meaning of ‘music’ in the complex auditory environment of the metro. Without the
ability to relate sounds in time and to organize them in meaningful ways (thanks in
particular to memory and expectation) he could not recognize music, nor could he perceive
the different events occurring in the auditory environment of the metro station. In this way,
we have replied to the first question addressed at the beginning of the present work: how is
the awareness of a musical event – a temporally extended object – constituted by the man
among the complex auditory environment of the metro?

Meyer also considers a musical work as a series of events that are apprehended
sequentially by a listener. An interesting observation is that “the apprehension of earlier
events has a marked effect on the apprehension of later events” (Levinson, 1997, p. 38).
Earlier events, in fact, in their relations with present events,\textsuperscript{122} influence the protentional
movement toward the future and the creation of expectations about what is going-to-be.
Along the same line, Levinson notes that “[i]n particular, the earlier portions of a piece,
through being heard, must invariably influence apprehension of later parts not contiguous
with them” (1997, p. 43). Every arriving event, therefore, in its relation to what has been,
generates expectations which may or may not be fulfilled, for events still later in the course
of the piece. Following music, listeners thus create moment-to-moment relationships that
are at the basis of their musical understanding. Following music moment-to-moment,

\textsuperscript{122} To make this point clear: “relationships between presently sounding events (accessed perceptually) and
earlier ones (accessed conceptually) are relations of a part to an earlier part, and not of a presently perceived
part to the whole” (Levinson, 1997, p. 476).
listeners become aware of what is going on. This process allows a synthetic and structural way of hearing music. In Stambaugh’s *Music as a Temporal Form* (1964), he asserts that “[s]ince the whole is never extended, it can and must be “contained” in each moment and in the passage from moment to moment. What is already past and what is to come, both are in the moment, not as objects remembered or anticipated, but as temporal functions essentially qualifying what is being heard. The ear cannot objectify as the eye does, and consequently its “synthesis” is of a different order. The ear has nothing in front of it; it cannot move from one thing to another and back again. Its synthesis is instantaneous and irreversible, a temporal synthesis” (1964, p. 273).

The concept of ‘basic musical understanding’ supports the idea that, in order to hear music with comprehension, ordinary listeners (those who do not have musical training) do not need particular musical knowledge. Perceptual understanding is distinct from an intellectual one in that it allows intuitive listening and immediate musical understanding. Levinson’s work is interesting in this respect: he suggests that ordinary listeners focus on moment-to-moment perception of music and on quasi-hearing; his book is a defence of the adequacy of intuitive, untutored listening. Thanks to structural unconscious processes involving the temporal structure, all listeners have the ability to understand a sequence of sounds as belonging to the same musical object.

I would like to return to the way in which a listener unconsciously creates, moment-to-moment, relationships between sounds and the place of expectation and anticipation in all of this. I argued for the idea that listeners have a crucial role in the organization of sound events, thus determining the meaning that a musical object and the related auditory experience have for a subject. The intentional tendency toward the future that characterizes the temporal structure and usually evokes an expectation is realized through a kind of chain reaction in which a present stimulus leads through a series of adjustments to a more or less specified consequent. Such a consequent is always implied in the tendency, once the tendency has been brought into play. Finally, when the consequent term arrives, the listener may be able to understand its relationships to the terms which have preceded it. Listeners therefore have a moment-to-moment response to the ongoing flow of music, during which
they constantly create expectations. Moreover, their auditory predictive schemas are updated for each new acoustic event in the sound environment.

Unyk and Carlsen (1987) also support the theory that listeners entertain certain expectancies about future events in music while musical events are unfolding. This way of relating sound events allows for the constitution of subjective musical experiences in the present perceptual act. Musical works are experienced as processes in progress that listeners grasp as the music unfolds. Indeed, listeners try to constitute meaning and create their own experience in the moment. From this perspective, musical listeners are like jazz improvisers: both are making up the music (or the musical meaning) as they go along, and they try to do it in a way that makes the most sense based on the preceding music. This also highlights the subjective character of musical hearing.

At this point it is necessary to better clarify a distinction: while retention-protention structure allows ‘listening in the moment’ and guarantees the possibility of having a perceptual musical understanding in hearing a sequence of sounds, the specific meaning (emotional or other) that the perception of a musical object acquires for a subject in the moment involves memory, expectation and anticipation, which differ from retention and protention because they are acts of consciousness referring to a content.

As I suggested previously, the expectations and anticipations (if any) that the man in the metro creates while he is perceiving the music that one or more musicians are playing somewhere in the metro influence, in different ways, his bodily and emotional responses as well as his beliefs. In this way, expectation and anticipation play an important role in determining the meaning that his perceptual experience acquires. A subjective musical experience is thus constituted. These observations reply to the second question I introduced at the beginning of the work: what are the main mental processes that determine the subjective character of the man’s musical experience in the metro?

Keeping the third question in mind – that is, what mental processes are involved in the specific case where the music the man in the metro hears music which is unfamiliar to him? –, we note that a listener also dynamically creates expectations as music unfolds during her first exposure to music or while listening to culturally unfamiliar melodies.
Studies have shown that, for example, listeners adapt quite well to the music of another culture after hearing it for only one minute, and listeners are able to recognize music during their first exposure to it\textsuperscript{123} – an ability that is often related to the process of expectation. In such studies, the participants exemplified listening to music \textit{in} the moment.

As McAdams and Bigand write, “to grasp this progression supposes that each of the new sound events can be connected to the preceding events. In cases where the incoming sound material is completely unknown to the listener, or does not correspond to known systems of structural relations (as would be found in the music of a given culture, for example), more primitive structuring procedures, such as segmentation into groups of events, may still operate directly on the relations among perceived surface attributes” (1993, p. 9). This suggests that listeners are able to create musical expectations and, thus, constitute meaningful musical experiences even when listening to unfamiliar music. The sense of misunderstanding that they might experience may be due to the incapacity to form anticipations. In the next chapter, I discuss this possibility by taking into account the perception of timbre.

\textsuperscript{123} See, for example, Marcus, Vijayan, Bandi Rao & Vishton (1999); Saffran, Aslin & Newport (1996); Saffran, Johnson, Richard & Newport (1999).
THE CASE OF TIMBRE EXPECTATION

Introduction.

As we have previously seen, the expectation process seems to play a very important role in the constitution of musical experiences, and it influences the subjective meaning they acquire. Most studies available until now – both in the field of music theory and the psychology of music – have analysed music expectation solely in relation to the melodic and rhythmic dimension of music, without paying much attention to other musical dimensions. By studying the when and what of something which is expected to occur during the unfolding of a musical event, empirical studies have investigated listeners’ ability to perceive and identify melodic and rhythmic patterns, as well as the influence that expectations and their violations have on this ability. In addition, they refer almost exclusively to Western tonal music and the expectations a listener creates on the basis of the musical structure which the composer has imposed on her musical piece.

The relevance of structural and hierarchical tonal relationships for the perceptual organization of music and the anticipation of the future ongoing of a musical piece is also supported by those philosophical theories which follow a formalist and architectonicist point of view in explaining the problem of music comprehension, or arguing for tonality as a ‘natural way’ of hearing music. Despite the great number of studies focusing on melodic and rhythmic expectations, less attention has been devoted to timbre expectations and their related emotional influence. As Gabrielsson and Lindström observe, “there is no systematic research on how the timbre of different musical instruments affects emotional

---

124 See, for example, Unyk & Carlsen (1987); Larson (1993; 2004); Lerdahl (2001); Margulis & Levine (2004); Margulis (2003); Aaeden (2003); Krumhansl (1997).
125 For example, Hazan, Brossier, Holonowicz, Herrera & Purwins (2007b); Hazan, Brossier, Marxer & Purwins (2007c).
126 As we have seen, Kivy (2001; 2002) among others.
127 In particular, Scruton (1997).
expression – certainly something that composers pay much attention to” (2001, p. 242). One reason for this lies in the fact that timbre is not the founding element of hierarchically organized systems as pitch is for harmony and melody. Moreover, it has acquired a greater role in contemporary music, which is not commonly the focus of theoretical and empirical research about expectation and its emotional influence.

In this chapter, I propose the perception of timbre as an example of ‘listening in the moment’ and as an additional phenomenon from which to explain the role of both expectation and anticipation. Particular attention is given to the case in which the man in the metro hears unfamiliar music. With reference to some contemporary music and music belonging to cultures that are very different from those most influenced by the Western tonal music tradition, I analyse some perceptual mechanisms that are used during musical listening and influence not only the organization of timbre, but also the way in which sound events are understood as music: the perception of similarities between sounds, the phenomenon of sound localization, and the recognition of the sound source. These three perceptual mechanisms allow for the creation of specific expectations based on timbre, and a musical comprehension of unfamiliar and unexpected sound events in the present moment of the perceptual act of listening.

The relevance of timbre in music.

Timbre involves a complex set of auditory attributes and it is processed by distinctive psychological activities. It is an important musical dimension for the organization of sounds during music listening. Moreover, it has a strong expressive power and an important role in the perceptual organization of other musical materials, such as pitch discrimination and the influence of the tonal context. As Andrew Earis and Patricia Holmes claim, “timbre is a fundamental and personal vehicle for conveying expressive intent, making a vital contribution to the communication of emotion in music, from performer to listener” (2007, p. 181). We can argue that a work’s instrumental colour often
makes a vital contribution to the structural features of a musical composition. It helps delineate form and can add expressive and depictive qualities that are central to the work’s character and identity. In this way, timbre is an essential part of what helps the man in the metro to define the identity and musical form of his various musical encounters.

Thus, timbre can be described as an expressive tool and a multidimensional attribute of sound, which allows a listener to distinguish one instrument from another. The most relevant features and attributes of sound that influence timbre processing are its brightness, roughness, richness, attack quality, the harmonics or overtones produced by the instrument’s vibrations, and the relative strength of the components of the frequency spectrum. These attributes are processed by specific auditory and cognitive processes and they have particular expressive power.

Some studies focusing on timbre which have presented convincing evidence that it can play an important role in music perception have been conducted at the Institut de Recherche et Coordination Acoustique/Musique (IRCAM). The goal of the IRCAM’s research is to determine – through the development of predictive models – the structure and components of the multidimensional perceptual representation of timbre. Additionally, this research is conducted in order to reply to two fundamental questions:
1.) what is the role of timbre in the creation and perceptual organization of musical materials?
2.) Does it play a role in the movement of tension and relaxation usually studied in relation to the melodic/harmonic and rhythmic/metric structure?

Other studies have explored the expressive properties of timbre, its nature and effects as providing a powerful channel for the direct communication of emotional and intellectual interpretative ideas.\textsuperscript{128}

\textsuperscript{128} The study The Role of Timbre in Expressive Musical Performance: A Case Study of Bach’s Prelude BWV 998 Played on the Acoustic Guitar by Earis and Holmes (2007) concerns the measurement of expressive variations in timbre. Manipulation of timbre is one of the main ways by which a performer communicates ideas and emotions. By studying the spectral qualities of the sound produced, individual differences in timbre were quantified. This type of analysis allows the objective study of expressive timbre. For further readings on the topic see Holmes (2010).
Because it is difficult to separate timbre from other perceptual variables, its study is often related to the study of its interaction with pitch discrimination or the perception of tonal structures (for example, the influence of tonal context on structural organization, the extraction of melodic structure, the perception of large-scale form and the sense of movement between tension and relaxation) and, therefore, to its influence on the perception and organization of other musical materials.\footnote{See, among others, Krumhansl & Iverson (1992); Warrier & Zatorre (2002); McAdams & Cunibile (1992); Deliege (1989); Margulis & Levine (2006). In a study by Margulis and Levine (2006), the identification of timbre pitches have been analyzed when they occurred in isolation, in relation to pitches of short melodies and with regard to the degree to which, given the preceding melody, pitches were expected. The authors also examined how timbre identification judgments were influenced by the melodic context and, subsequently, how melodic expectations were affected. The results showed the improvement of timbre identification in relation to the degree to which the pitches were expected, supporting the idea of an interaction between timbre and pitch. This means that timbre is often processed together with other musical elements, which could explain why it cannot be at the basis of a structural organization of music.} In Lerdahl and Jackendoff’s view, for example, timbre is a dimension of musical structure which is not hierarchical in nature. It plays an important role in their \textit{Generative Theory} in that it makes crucial contributions to the principles that establish the hierarchical structure of a piece of music (1983a, p. 9). In this way, timbre helps the perceptual organization of sounds into musical structures.

Particularly for contemporary composers, timbre may be considered as an important perceptual cue and a form-bearing dimension\footnote{Form-bearing dimension is a concept proposed by McAdams (1989) which designates the perceptual cues used by listeners to evaluate the similarity between musical materials and derive a coherent perception of the musical structure.} which contributes to the organization of music. In this case, for example, psychoacoustic attributes of timbre such as roughness are important in creating the movement between tension and relaxation in non-tonal sonorities (McAdams, Vieillard, Houix & Reynolds, 2004). In Western tonal music, which is highly hierarchically structured, timbre has not been considered as a basis for musical systems. In spite of its great contribution to the structural organization of sounds (particularly in music which is very rich in timbre, like the case of Björk’s songs), music perception is usually difficult to conceive only on the basis of timbre organization.

Studies such as those presented by Krumhansl and Iverson (1992) suggest that timbre is perceived more in absolute than in relative terms. Studying whether timbre is perceived in interaction with other attributes of musical tones (such as pitch), results
showed that the relative perception of timbre is considerably weaker. Based on this difficulty in perceiving relationships based on timbre alone, the authors conclude that “it seems unlikely that attempts to establish hierarchically organized patterns of timbres (Lerdahl, 1987) or to effect transformations of timbre patterns such as transposition and inversion (Slawson, 1985) will be successful” (1992, p. 750).

This creates some problems for the generation of musical expectations based on timbre. Expectations are usually conceived as arising within a structural tonal context. If timbre cannot form the basis for a musical structural system, it follows that – according to these theories about expectation – timbre alone cannot cause expectations to arise. But what about the case of music which is not based on a tonal system and assigns an important role to the manipulation of the acoustic features of sound, such as some contemporary compositions by Debussy, Varèse and Schöenberg, or electroacoustic music?

The creation of expectations requires the establishment of relationships between sounds. The majority of studies about musical expectations presuppose the perception of tonal and melodic structures as the condition for both music comprehension and emotional responses to music. Can timbre alone bring about expectations and emotional responses without recurring to tonal, established musical structures? This question has been addressed by Lerdhal in his work *Timbral Hierarchies* (1987): is it possible to build up hierarchical mental representations based on timbral relations in tone sequences? By referring to other literature focused on a similar question and following Schöenberg’s desideratum for a timbre-based music, he proposes that timbre can be organized hierarchically and, by applying two hierarchical structures (grouping and prolongational), he suggests the possibility of creating timbral spaces which move in musically meaningful patterns.

Even if it has traditionally functioned associationally, in some cases (as in contemporary music and works like Messiaen’s *Chronochromie*, 1960) timbre can be the principal bearer of the musical form and, from an aesthetic point of view, it can reinforce

---

[131] As we have seen, Meyer (1956); Narmour (1990).

[132] See Ehresman & Wessel (1978); Risset & Wessel (1982); Zhang & Ras (2007). Ehresman and Wessel, for example, claim that timbres can be perceptually related to each other and that “timbre can be related and even transposed in a manner analogous to pitch”. This is fundamental for a structural organization of sounds.
the pitch structure. As discussed by Schaeffer (1966), timbre has to do not only with the acoustic features of sound, but also with the form of music.

Another way to face the problem consists of making a parallel with language on the basis of the growing evidence that speech and music share mechanisms for sound category learning and that the processes which create sound categories have a significant degree of overlap. In the first chapter of his work *Music, Language, and the Brain* (2008), Aniruddh Patel compares music and speech in terms of the way people organize pitch and timbre. He also explores why timbral contrasts are rarely the basis for musical sound system. He observes that two common properties given particular importance in human music include an organized system of pitch contrasts and musical timbre:

from an aesthetic standpoint, timbre is arguably as important as pitch as a perceptual feature of music. (Imagine, for example, the difference in the aesthetic and emotional impact of a jazz ballad expertly played on a real saxophone vs. on a cheap keyboard synthesizer.) From a cognitive standpoint however, timbre differs sharply from pitch in that the former is rarely the basis for organized sound contrasts produced by individual instruments. Of course, timbral contrasts between instruments are quite organized, and are used in systematic ways by composers from numerous cultures, for example, in Western, Javanese, and African instrument ensembles. A notable study in this regard is that of Cogan (1984), who conducted spectral analyses of symphonic music and proposed a theory of timbral contrasts inspired by linguistic phonetics (Cogan & Escot, 1976). However, the salient point is that organized systems of timbral contrasts within instruments of a culture are rare. Why is this so? (2008, p. 28).

Why do timbral contrasts rarely serve as the basis for musical sound systems? Some authors have tried to reply to this problem by underlining that there are both physical and cognitive reasons for the rare use of timbre as the basis for organized sound contrasts in music and for the creation of musical scales and intervals. A few notable exceptions are represented by instruments like the simple Jew’s harp, which emphasizes rapid timbral contrasts, and the Australian didgeridoo, which is renowned as an instrument in which timbre is the primary principle of sonic organization.

---

133 See, for example, Krumhansl (1989); Ehresman & Wessel (1978); McAdams & Cunibile (1992). Ehresman and Wessel (1978) and McAdams and Cunibile (1992) have investigated whether listeners are capable of hearing ‘timbre intervals’ through research on perception of timbral similarities between instruments. In Patel’s opinion, “their research raises the possibility that timbre relations are not perceived with enough uniformity to provide a basis for a shared category system among composers and listeners” (2008, p. 34).
Some examples of music based on timbral contrasts (‘timbre-based musical systems’) have been highlighted by Patel:

this example [North Indian Tabla drumming] illustrates how a successful “timbre music” can be built from a speech-like system of organization. Another long-lived musical tradition which emphasizes timbral contrast is Tibetan tantric chanting. In this music, timbral contrasts slowly unfold while pitch maintains a drone-like pattern. A key to the aesthetic success of this tradition may be that it takes an aspect of human experience that is normally experienced extremely rapidly (timbral speech contrast) and slows it down to a completely different timescale, thus giving the opportunity to experience something familiar a new way (2008, p. 53).

Making a parallel with language, Patel notices that there are other ways to organize timbral systems, from which Tibetan tantric chanting might have been inspired. Unlike music, the primary dimension for organized sound contrasts in language is timbre; the human voice is the supreme instrument of timbral contrasts.

Things have changed with non-notated contemporary music, which has opened up endless timbral possibilities and increased the range of possible sound organization. In this case, there is a growing timbral common praxis and a widespread acceptance of timbre as a primary surface feature and as a structural device in music across genres (Rudy, 2007). Since the 20th century, timbre has become an important factor in music composition. Starting with musique concrète (1948), electronic music (1950) and digital music (1957), timbre not only indicated the source of sound, but it began to also qualify the form and structure of music, becoming an important tool for new kinds of musical compositions in Western culture. However, no effective timbre theory for music composition has been developed as of yet.

A problem with this type of timbrally-qualified music (which also involves computer, electroacoustic, atonal and improvised music) is that it remains unfamiliar to the majority of Western listeners. Because of their ‘tonal’ exposure, most find it quite difficult to listen to such music and to organize its sounds meaningfully. This difficulty influences appreciation and comprehension, but does not necessarily mean that contemporary music lacks a strong expressive and emotive potential for untrained listeners. As we have previously seen, even when music is unfamiliar, listeners are able to create expectations
based on their knowledge and past experiences: the fulfilment or violation of such expectations may result in emotional responses.

**Timbre and expectation.**

Because of the relevance of timbre in music perception, I will next explore some of the ways in which timbre might give rise to expectations, particularly when people listen to unfamiliar contemporary music.

Expectations are usually thought to be created in relation to the structural features of music and the various probabilities of their occurrences (i.e. expectations are built on the basis of the most recurrent musical structures in a given environment). According to different theories and studies,\footnote{For example, Meyer (1956); Juslin & Västfjäll (2008); Lerdahl & Jackendoff (1983a).} musical expectations involve syntactical relationships between different parts of the musical structure. As we have seen, timbral relations are not at the basis of musical systems, which is one of the reasons why it is difficult to study timbre expectations and their influence on the musical experience.

Studies about the creation of expectations in relation to the timbral dimension of music have only recently been developed. A few such studies have investigated the kind of timbre that is expected, for example, at the end of a sequence (Warrier & Zatorre, 2002). Others have studied the expectation of timbre categories through a computational model of \textit{what/when} expectation (Hazan et al., 2007c). Some researchers have also investigated timbral motion and timbral crescendo in contemporary music, both of which create expectations and have a powerful emotional impact (McAdams, Vieillard, Houix & Reynolds, 2004). Finally, and based on this last point, several studies have been developed in order to analyse how changes in timbre influence listeners’ emotional responses.\footnote{See, for example, Padova, Santoboni & Olivetti Belardinelli (2005); Padova, Blanchini, Lupone & Olivetti Belardinelli (2002).} They give evidence that the expectations determined by the timbral features of sound represent a
particular auditory experience eliciting specific emotions. In addition to tonality and meter, these studies assert that timbre may be interpreted as the basis for the constitution of a class of auditory schema whose predictability create opportunities for both the pleasure arising from accurate prediction and the contrastive valence arising from innocuous surprises. This concept – that is, that a listener creates associations in relation to timbral features – is more viable in particular kinds of music, such as some of the contemporary styles discussed.

Listeners create timbre expectations by searching for regularities in the auditory environment and in relation to the acoustic features of sounds. In the case of unfamiliar music which is not tonally structured, timbre becomes a fundamental cue for the organization of sounds and the creation of expectations. This is clear in studies that consider the influence of timbre in the musical perception of both musicians and non-musicians. The fact that non-musician subjects organize music on the basis of timbre shows its relevance for the perception of unfamiliar music.

In the context and for the purposes of the present work, I focus on the relevance of timbre in the perception of unfamiliar contemporary music. By using the description ‘unfamiliar contemporary music’ I am describing all those contemporary musical genres in which tonality has been replaced by other organizational systems, such as serial music, electroacoustic music, concrete music, experimental music, atonal music, and minimalist music; and those for which the sound experimentation has a great relevance, as in the case of post-1975 contemporary classical music. Specific examples of post-1975 contemporary classical subcategories, include post-modern music, spectral music, post-minimalism, and sound art. Because of their minor diffusion with respect to musical genres based on tonality, these contemporary musical genres are relatively unknown to the majority of non-musician listeners who are accustomed to Western tonal music. For this reason, unfamiliar contemporary music compositions are more difficult for untrained listeners to appreciate.

If, according to Meyer, the principal source for music’s emotive power lies in the realm of expectation, and if it is true that timbre is able to arouse particular expectations, then timbre also arouses particular affective and meaningful experiences in a non causal way.

See, for example, Pitt (1994); Paquette & Peretz (1997); Poulin-Charronnat, Bigand, Lalitte, Madurell, Vieillard & McAdams (2004).
and understand. The same issue arises for music belonging to cultures which are musically very different from that of the listeners. In these cases, listeners find it difficult to organize sounds into musical structures. A common thought is that they are not able to create correct expectations because they are unfamiliar with that music, and this leads to a sense of misunderstanding.\textsuperscript{138} I argue for the hypothesis that, even in these cases, listeners are able to create expectations and to musically understand their auditory experience thanks to a sound organization which is made \textit{in} time and on the basis of the timbral features of music.

As we have seen in the fifth chapter, expectations are usually related to the organization of musical sounds into structures, but they do not always reflect the way music is actually organized or has been structured by the composer. Expectations are firstly influenced by the listeners’ musical knowledge and experience, which does not always correspond to the music they are actually listening to. This is clear in the case of unfamiliar music. Secondly, listeners are not perfect in their musical expectations and, when they anticipate sounds, they often rely on heuristic knowledge. It follows that they do not always grasp the musical organization, particularly in those cases in which music is organized in new ways and on the basis of other perceptual cues. This does not necessarily mean that they are not able to meaningfully organize sounds, to appreciate music, to create expectations and have emotional experiences. They simply find new dynamic ways to organize it.

In the case where listeners have always had exclusive exposure to Western tonal music, the process of listening to music which is organized differently becomes more difficult; as, for example, in the case of unfamiliar atonal contemporary music. The principal reason for this difficulty is that during their ‘tonal’ experience, listeners unconsciously acquire tonal schemas that they use to organize the music they listen to and to create expectations about the future. In the case of unfamiliar atonal music, such

\textsuperscript{138} Theories concerning emotional experience in music are often related to the question of musical meaning. Following Meyer (1956), emotions could be considered as one of the possible meanings that a musical experience can acquire and they depend upon the creation of expectations: the way they are satisfied, delayed or blocked.

A common idea is that if a stimulus does not arouse expectations, it is meaningless. According to this idea, music in a style with which listeners are totally unfamiliar is meaningless because the listener is not able to create expectations.
expectations are not satisfied and the listener is not able to create correct expectations about
the future ongoing of the music, nor are they able to anticipate it. Listeners therefore have
to find new ways of meaningfully organizing sounds and anticipating the future. One
musical dimension that could help them to accomplish this is timbre: in relation to the
timbral features of sound listeners try to find regularities and recognize musical passages. It
follows that the listener creates expectations in relation to the timbral regularities that are
internalized during the temporal unfolding of a musical event and its ‘listening in the
moment’. Thus, they acquire the ability to follow and involve themselves in the music
during the present act of listening to it.

This has consequences for the constitution and subjective meaning that the listener’s
perceptual experience acquires while hearing music. If we follow Meyers’ theory, listeners
feel emotions when their expectations are broken or not fulfilled. I argue for the idea that,
while listening to unfamiliar music in which timbre has a strong expressive power, listeners
create expectations during the temporal development of music and on the basis of the
relationships that they create between sounds. If such expectations are not realized or if
they influence the perception of other musical materials, the listener constitutes different
musical experiences and therefore may have different emotional responses. For this reason,
I propose timbre expectation as an important mental mechanism for listeners’ emotional
response to unfamiliar contemporary music.\textsuperscript{139}

One argument supporting the relevance of timbre in the perception of unfamiliar
music comes from the study \textit{Perception of Pitch and Timbre by Musically Trained and
Untrained Listeners} by Mark Pitt (1994). By comparing musicians and non-musicians on
pitch and timbre perception, Pitt showed that non-musicians’ judgments were more affected
by timbre differences than those of the musicians. As Paquette and Peretz observe, “for
nonmusicians, timbre is probably a more salient dimension than pitch because the former is
generally more informative about environmental events and these listeners have not been
trained to analyse pitch closely” (1997, p. 695). Non-musicians are therefore particularly

\textsuperscript{139} Timbre is often an important perceptual cue and a form-bearing dimension that allows a dynamic and
immediate organization of sounds into musically meaningful forms, as well as a real-time creation of
expectations. In the case of new instruments, for example, listeners easily learn to recognize and organize
timbral features of sound (Paquette & Peretz, 1997).
sensitive to timbre factors and this is an important ability for the perception of contemporary unfamiliar music.\textsuperscript{140} Such observations also show that people do not have an innate tendency to hear music tonally because different ways of hearing music exist.

An important aspect of expectations is that they are built dynamically through experience and thanks to a process of learning. The case of timbre perception in listening to unfamiliar music is particularly interesting because it shows an ability to find regularities and learn how to relate sounds in the present moment of the act of listening. It also indicates that listeners develop different ways of hearing and perceiving music, creating expectations that do not necessarily coincide with the unfolding of the notated or established musical structure. The aim of the next paragraph is to analyze some of the processes that allow for the creation of expectations based on timbre and determine the extent to which timbre influences the sound organization of unfamiliar musical events.

### Cognitive and perceptual processes involved in the organization of timbre.

A problem which is related to my proposal concerns the comprehension of the way listeners create expectations based on timbre. What are the mental processes and cognitive principles involved during the organization and processing of the timbral features of sound? How does timbre influence the expectations and musical experience of listeners who listen for the first time to unfamiliar contemporary music?

In general, we can understand how listeners build expectations by studying the mental mechanisms and processes of organization of the sound environment that are involved in music perception. Some of these are well explained by Gestalt psychology (grouping principles), while for the most part, such mechanisms and processes can be best examined through reference to empirical research studies which focus on the way listeners learn the regularities of the music to which they are exposed.

\textsuperscript{140} See also Poulin-Charronnat, Bigand, Lalitte, Madurell, Vieillard & McAdams (2004).
I focus my attention on three cognitive processes that are particularly relevant for the creation of timbre expectations while hearing unfamiliar music: the identification of the sound source, the perception of timbre similarities and the perceptual phenomenon of sound localization.

Recognition – defined by Clarke (2005) as that kind of perception for which the perceptual system has become adapted (or tuned) – is very important for everyday listening. It involves detecting the objects and events in the world that are specified by sounds. This is what the man in our example does in order to orient himself in the complex auditory environment of the metro station and respond to the information he receives (which, in this way, acquires a specific meaning for the man in his context and situation).

Sound is usually perceived as referring to something (the guitar being played somewhere in the metro, the woman who is calling her children, the loudspeaker announcing that the metro is closing) and timbre identification is important for the recognition of each sound source. A sound source usually has a particular meaning for the subject, which drives her to create specific expectations. For example, recognizable sources create expectations about what ‘might’ come next (Rudy, 2007): their fulfillment or violation may arouse emotions. The attribution of the cause of the sound is thus important for the effect it has on the listener’s musical experience. Different instruments, for example, have different meanings, affective properties and emotional effects on a subject.

Studies have shown a significant interaction between instrument and emotion judgment and that timbre identification affects the perception of emotion in music, even in cases of unfamiliar music and musically inexperienced listeners.\(^{141}\) The perception of emotion conveyed by a melody can be affected by the identity or timbre of the musical instrument on which it is played. This is one example showing the way timbre perception may bring about expectations with emotional effects: the relationship a listener creates between the sound and its source makes him expect certain timbral qualities or events which, when realized or violated, may have emotional effects. Such emotions characterize

\(^{141}\) See, for example, Hailstone, Omar, Henley, Frost, Kenward & Warren (2009).
one of the possible subjective meanings that the listener’s musical experience might acquire.

In the case of some contemporary music, things are more complex for the listener because of the use of ‘anything’ as an instrument. The introduction of noise in music, for example, has created uncertainty. This has had a particular influence on the arousal of expectations and the discovery of new sound information. Ambiguity, in fact, is an interesting aspect of expectations which gives rise to particularly strong tensions and powerful emotions. Because of ambiguity, listeners are not able to relate sounds to a specific source, or they relate them to a sound source that may normally have another function – such as the scraping of a bow on a violin or the sound of a drill. At the same time, this opens up the possibility of new subjective interpretations and associations which carry different emotional experiences.

The introduction of the sound of the train in Étude aux chemin de fer (1948) by Pierre Schaeffer, for example, mimics the sound of an object that has extra-musical meaning – that is, a train – within a musical context. Such a meaning is related to the listener’s daily experience in the world. In a musical context, the sound of the train does not ‘act’ in its usual way: the function here is different. This creates uncertainty and determines the generation of different kinds of expectations.

In order to study the influence of timbre, we might take as an example the case of acousmatic music,¹⁴² which is based on the effect of unseen sound sources and may contain

---

¹⁴² ‘Acousmatic music’ is a form of electroacoustic music that deals specifically with acousmatic sound as a compositional resource. ‘Acousmatic sound’ is sound one hears without seeing its source or cause. The term acousmatique was first used by the French composer and pioneer of musique concrète, Pierre Schaeffer (1966). In acousmatic art one hears sound from behind a ‘veil’ of loudspeakers, the source cause remaining unseen. Schaeffer held that the acousmatic listening experience was one which reduced sounds to the field of hearing alone.

This musical practice has a historical basis in musique concrète, through which music began to be created using non-acoustic technology, and existed only in a recorded format intended for reception via loudspeakers. The compositional material of such genres is not restricted to the inclusion of sonorities derived from musical instruments or voices, nor to elements traditionally thought of as ‘musical’ (melody, harmony, rhythm, meter and so on), but rather admits any sound, whether acoustic or synthetic. With the aid of various technologies, such as digital signal processing tools and digital audio workstations, this material may then be combined, juxtaposed, and transformed in any conceivable manner.

In an acousmatic concert the sound component is produced using pre-recorded media, or generated in real-time using a computer. The sound material will then be distributed spatially via multiple loudspeakers, which is a practice known as diffusion.
sounds that have recognizably musical sources, but may equally present recognizable
sources that are beyond the bounds of traditional vocal and instrumental technology. The
techniques of synthesis and sound processing that are employed may present us with
sounds that are unfamiliar and that may defy clear source attribution. In the case of this
kind of music, timbre is very important in order to distinguish and recognize sound sources
because it makes us expect another event with a particular timbre and a particular meaning.

When considering the way listeners experience new music, a crucial topic in the
field of music psychology is the perception of similarity between musical materials. As we
have seen in the fifth chapter, such perception influences the organization of sounds and,
consequently, the perceptual understanding of events occurring in the auditory
environment. In the case of timbre perception and listening to unfamiliar music, similarity
is particularly relevant because it does not refer to the concept of hierarchy, but to that of
association – i.e. the relative closeness or distance that is perceived to exist between
musical objects or ideas. Thus, even if the listener is not familiar with the musical style and
its formal structures, she is able to meaningfully organize the sounds constituting music.

Studies in the fields of music psychology, cognitive science and neuroscience have
investigated how similarity relations among new musical materials are apprehended by
listeners, what the criteria underlying listeners’ perception of similarities are and the degree
of similarity between timbre representations. An assumption of some of these studies is
that listeners use similarity principles to establish relations among musical materials in
order to organize them into classes. In the case of timbre, similarity groupings are formed
by association of surface resemblances, and not by hierarchical relationships. Whilst
listening to non-tonal music, for example, listeners encounter a large variety of unfamiliar
musical materials. In such cases, pattern processing cannot be entirely based on previous
schematic knowledge and listeners have to learn to use other cues. The cognitive system
has to adapt to new uses of musical materials in order to perceive the relations among them.
Without previous schematic knowledge, similarity perception is used to organize sound
events into musical structures.

143 See McAdams, Vieillard, Houix & Reynolds (2004); Toiviainen, Tervaniemi, Louhivuori, Saher,
Moreover, the structural hierarchy of atonal music can be inferred through the relative salience of events and musical elements such as timbre. Timbre is thus considered, together with its associated attributes, as a form-bearing dimension. Such organization determines the integration of non-tonal materials into a musical structure and the creation of expectations which could have emotive effects. In this case, the perception of relationships between sounds is more associative than hierarchical (Dibben, 1994) and largely developed through similarity principles. The perception of similarity between musical materials thus accounts for a large part of the listener’s musical experience in cases of unfamiliar musical styles. In such cases, the listener’s perception of associations, formation of musical categories and their sense of familiarity are all possible because similarity principles allow for the comprehension of musical relationships in contemporary music independent of musical training.

Sound localization is the sense of spatial position of a sound source. Knowing where music comes from is an important part of knowing one’s environment and responding to it adequately. Moreover, it is a perceptual phenomenon that strongly influences the listener’s expectations. From a general perspective, the world is full of many different sounds of various elevations and timbres at any given time, and spatial location is an innate mechanism important to one’s ability to distinguish and understand their sound environment. It also has a biological meaning: the ability to detect a sound’s source in dangerous situations is important for survival. In the case of music, it reflects a natural tendency which is influenced by experience and learning.

Changes in sound locations involve where expectations. Musicians have sometimes manipulated the locations of sounds (as in the antiphonal works of Giovanni Gabrieli or the electroacoustic works of Karlheinz Stockhausen), but they have less often manipulated listener’s expectations of location.\(^\text{144}\) An interesting aspect of timbre is related to the tendency to perceptually connect sound events arising from the same sound source: successive events that are relatively similar in their spectrotemporal properties may be perceived as arising from the same source and thus tend to be grouped together (McAdams,

\[^{144}\text{See, for example, Bernard (1983); Trochimczyk (2001).}\]
Vieillard, Houix & Reynolds, 2004). Timbre perception is independent but related to the phenomenon of sound localization, and timbre contributes to the creation of expectations about where a sound source is located. Because of the relevance of timbre to the identification of sound source, changes in the location of a sound source may affect the perceptual organization of timbral features.

**A proposal for further empirical studies.**

In order to better understand these processes and the influence of timbre-based expectations on listening to unfamiliar musical events, I suggest the study of some specific situations where Western ‘tonal’ listeners are involved. In particular, I propose the analysis of two contexts and situations where similarity and sound localization play a specific role in creating expectations based on the timbral features of music:

1. live contexts where performers improvise on the timbral features of a musical piece;
2. live contexts where performers change the location of the sound source.

The basic idea is that the interaction between the processing of acoustic features of sound, cognitive mechanisms like sound localization and similarity, and the temporal structure of human consciousness, allows the listener to create expectations about the ongoing of the musical event even in the case where she is not familiar with it and something new is experienced. The aim is to discuss the role of timbre in such a process.

During their experience with unfamiliar sound events, listeners learn to relate sounds on the basis of similarity relations and associations that they find between the timbral features of a musical event. Such relations allow them to create expectations about the future ongoing of music (the what, when and where specific timbral features or events are expected to occur). In improvised music, it often happens that listeners encounter unexpected sound events that create a sense of surprise. These sound events break or violate the listeners’ expectations and thus contribute to the constitution of a particular musical experience which might have, for example, a particular emotional content. Surprise
is a cognitively ‘simple’ emotion; a feeling which by its nature has – or can have – a relatively simple cognitive content.

Although in Huron’s opinion surprise represents a failure from the biological point of view, in this case it guarantees a dynamic experience and involvement with music, as well as a musical experience which is emotionally meaningful. Timbre is a musical dimension that strongly contributes to the feeling of surprise. In the case where musicians improvise on the timbral features of a musical composition, the expectations a listener dynamically creates – thanks to the immediate learning of timbre regularities and their similarities – are often broken. The listener thus experiences a state of surprise, which is a common response to unexpected stimuli and may be followed by other emotions such as fear, joy, sadness, etc. This is particularly common in those cases where listeners experience unfamiliar music for which timbre plays a relevant role. In this way, timbre perception strongly influences the creation of expectations and the following emotional experience. At the same time, the timbral features of sound are used by the perceiver for an immediate and dynamic organization of the new sound events.

I have previously proposed the possibility of a kind of interaction between timbre perception and the phenomenon of sound localization in the creation of expectations. In cases where musicians change the location of the sound source, people never know exactly where the sound comes from. In this way, they create new unexpected events and a sense of surprise. The listener perceives a spatial movement which creates uncertainty and may cause the breaking-off of her expectations.

In the case of sounds coming from unexpected locations, there is an increase of information which requires more attention and dynamic behaviour on the part of the listener in order that they may organize and reply to the new events. This ability to appropriately attend and adapt to sounds coming from unexpected locations determines the creation of new expectations and the ability to find regularities between sounds. Changing the location of the sound source may affect the listeners’ expectations, thus determining new emotional experiences and a different comprehension of music.
To conclude, in the case of unfamiliar music or the first listening to an unknown musical piece, it seems that listeners are able to create expectations on the basis of the timbral features of sound thanks to the perceptual phenomenon of similarity and sound localization. The ability to follow timbral information in this way facilitates an immediate comprehension of the sound environment, which may subsequently lead to the experience of pleasant emotions.

Some relevant features of the expectation process.

The aim of this paragraph is to highlight some relevant features of the expectation process that emerge in the case of timbre perception and in listening to unfamiliar music. They are particularly relevant for the constitution of musical experiences in the moment and in complex auditory environments where musical events are not expected to occur.

Through the previously discussed situations concerning the perception of timbre, it has been noted that listeners learn to create expectations dynamically during the temporal development of music, particularly when they encounter unfamiliar musical events. This dynamic ability may be, for example, due to the processing of similarity relations based on timbre. Following Huron’s terminology, in these cases listeners create dynamic expectations which arise ‘on the fly’ and are linked to short-term memory (2006, p. 227-231). Such expectations are the result of short-term patterns that are updated in real time. 

145 ‘Short-term memory’ refers to the capacity for holding a small amount of information in mind in an active, readily available state for a short period of time. It is defined as a system for temporarily storing and managing information required to carry out complex cognitive tasks such as learning, reasoning, and comprehension. In the case of music listening, short-term memory is, according to numerous laboratory-based studies, capable of storing sound sequences for up to five seconds in duration. Such tests have also frequently examined participants’ ‘memory span’, which is defined as the number of items, usually words or numbers, that a person can hold onto and recall in short-term memory (see, for an example, Huron, 2006, p. 227). One of the keys to retaining a pattern in short-term memory is repetition and, in the case of music, there is plenty of repetition to help cement patterns into memory. Thus, short-term memory is particularly relevant in the case a subject listens to unknown music: it helps her to organize sounds and create expectations in the moment. Short-term memory is related to working memory. Such a relationship is described differently by various theories, but it is generally acknowledged that the two concepts are distinct. Working memory is a theoretical framework that refers to structures and processes used for manipulating information. Short-term memory
They are shaped by immediate experience, such as when exposure to a novel work causes a listener to expect similar passages as the work continues. The fulfilment or violations of these dynamic expectations have an emotional effect on the listener’s musical experience.

Some examples of experiments supporting the idea that listeners dynamically create expectations as music unfolds and during the first exposure to music are those of Jenny Saffran (1996; 1999) and Gary Marcus (1999), both of which have been conducted with infants. In the case of culturally unfamiliar melodies, studies conducted by Tuomas Eerola (2004) and Mary Castellano et al. (1984) showed that listeners adapted quite well to the music of another culture after hearing it for only one minute. These studies were used to explain dynamic expectations arising from immediately preceding musical experiences. To accomplish this task, they used continuous response methodology which allowed the researchers to study moment-to-moment variations in the listeners’ responses without interrupting the flow of the music.

As I have suggested, when ‘tonal listeners’ listen to unfamiliar contemporary music (lacking a strong tonal structure), they are still able to find ways of organizing the music so that they may be involved in its temporal development and have a meaningful experience. The perceptual organization of timbral features, as well as the creation of expectations during the temporal unfolding of the musical experience may both occur as to fulfil the need for a kind of comprehension and appreciation of unfamiliar music. In these cases, listeners are able to dynamically organize sounds into musical structures, even if they do not know the structure of the music that they are hearing and if the perceived organization does not correspond to an established structure. Sensitivity to the timbral features of music and the creation of relationships based on similarity is important for the creation of expectations, which ultimately affect the subjective character and meaning of the musical experience.

generally refers to the short-term storage of information, and it does not entail the manipulation or organization of material held in memory. Thus while there are short-term memory components to working memory models, the concept of short-term memory is distinct from these more hypothetical concepts. Working memory is the executive and attentional aspect of short-term memory involved in the interim integration, processing, disposal, and retrieval of information.
The ability to form dynamic expectations based on timbral features also indicates that particular attention is paid to the regularities of the environment in general, which forms the basis for the process of statistical learning – i.e., learning how frequently a particular event occurs (Huron, 2006). The ability to respond to unexpected or unfamiliar events and to find new cues for the organization of sound shows a process of learning that is based in the present moment of a musical experience. Such a process may result from simple exposure, without necessarily relying on explicit processes of analysis.\textsuperscript{146} It is part of the dynamic relationship that human beings have with the environment in which they live and are in constant communication with (Merleau-Ponty, 1945). In this way, listeners constitute – intentionally form objectivity by virtue of their subjective encounter with the environment (Sokolowki, 1964a) – subjective musical experiences which can be emotionally meaningful. Emotion is here understood as a meaning which the musical experience may acquire; a meaning that is constituted during the musical experience as it progresses.

As we have seen, another important factor that contributes to the organization of sound sequences into musical forms is the temporal structure of human consciousness. This is evident when we consider the definition of expectation as a basic strategy of the human mind that is able to bring past experiences to bear on the future. In the case of musical expectations generated in relation to the timbral features of sound and during the temporal development of unknown music, short-term memory plays an interesting role. As indicated by some cognitive studies, listeners tend to internalize most musical structures in long-term memory, but in the case of timbre it is short-term memory which contributes to the creation of dynamic timbral expectations. On one side, in this way timbre gives something fresh to the perception of the same musical phenomenon and it may explain why listeners continue to be interested in music they have heard more than once. On the other side, short-term

\textsuperscript{146} Tillmann and McAdams (2004) observed that “becoming sensitive to the statistical regularities in the timbre sequence might result from simple exposure without necessarily relying on explicit process of analysis” (2004, p. 1138).
Moreover, “if we consider timbre to be one of the main perceptual vehicles for sound source identity, studies demonstrating implicit learning sequential relations among timbres may shed light on the acquisition of knowledge about sequential expectancies and dependencies among sources in the natural environment” (2004, p. 1139).
memory allows for the creation of timbre relations *during* the process of listening to unfamiliar musical works. Therefore, short-term memory seems to be very important in enabling a listener to encode and maintain musical similarities during the unfolding of a musical piece and in a dynamic context. This allows the listener to find relationships and structures between sounds, to create expectations while listening to new musical materials and, finally, to have a meaningful musical experience.

To conclude, short-term memory and the temporal structure of human consciousness are strictly related to the dynamic aspect of human perception. A musical phenomenon is a dynamic succession of events, and listeners create expectations as music unfolds. The human mind must be open and ready to respond to new and unknown events occurring in the auditory environment. As previously discussed, in the case of contemporary unfamiliar music, sounds are organized on the basis of timbre regularities during the unfolding of the musical event. This allows for the creation of different kinds of expectations (regarding timbre, the identity of the sound source and its spatial location), whose role in the perception of unfamiliar music is a clear example of ‘listening in the moment’. This ability to establish new and dynamic expectations indicates the way listeners constitute their musical experiences *following* music, and *during* the very act of listening to a musical event. The temporal structure of human consciousness finds here its exemplification and it becomes clear how important the ability to create relationships between sounds *in* time is. Without this ability the listener would not be able to meaningfully relate sounds and, consequently, appreciate music. The sense of misunderstanding coming from a lack of sound organization negatively impacts the possibility of one’s musical enjoyment.

**The case of the man in the metro.**

The main idea that I tried to develop in this chapter is that ‘tonal listeners’ find a way to organize sounds into musical forms, to be involved in the temporal development of
music and to constitute musical experience even when they listen to unfamiliar musical styles which lack a strong tonal structure. Perceptual organization of the timbral features of sound and the creation of expectations during the temporal unfolding of a musical experience allow for a kind of comprehension and appreciation of unfamiliar music. In these cases, listeners are able to dynamically organize sounds into meaningful musical structures even if they do not know the structure of the music they are hearing and if the perceived organization does not correspond to an established structure.

As pointed out by Rotheberg (1996), the experience of new musical technologies that synthesize sounds in different or unfamiliar ways, for example, suggests new combinations and rearrangements of what is familiar in the world, as well as sound relations which cannot be judged or evaluated based on any familiar criteria. For some listeners this means that musical works taking advantage of these new technologies cannot be appreciated, while for others it is the beginning of the freedom of the new.

The creation of timbre-based expectations is important to the organization of sounds and for the meaning that the listener’s subjective experience acquires. Moreover, it is important to be able to react to new and unexpected auditory events and to constitute musical meanings in the case of unusual sound experiences.

The main question I introduced at the beginning of the work considered how awareness of a musical event might be constituted by a man in the complex auditory environment of the metro station. Based on the information presented in this chapter, we may now conclude that the case of timbre-based expectations and the way listeners organize unfamiliar music during its temporal development exemplify the process of ‘listening in the moment’ which characterizes the man’s experience in the metro and allows him to constitute and be immediately aware of his musical experience. In this context, timbre also plays an important role in his perception of music: it enables him to distinguish musical sounds from other sound events occurring in the auditory environment, and to organize certain acoustic events into musical forms even if the music is unfamiliar or completely unexpected in that context. Some of the processes that we have previously taken
into account play a relevant role here. In particular these include similarity, sound localization, and recognition of the sound source.

When the man in the metro encounters an unfamiliar musical event, he is thus able to understand it as a *musical* event thanks to timbre organization. The structure he perceives likely does not correspond to the notated one, but timbre organization and the expectations based on such organization allow him to constitute a musical experience. Timbre is therefore an important musical dimension for the man in the metro.

In relation to my second question – what are the main mental processes that determine the subjective character of the man’s experience –, expectations and anticipations are very important. Both of these processes are particularly relevant for the perceptual comprehension of music in everyday situations, which include encounters with unfamiliar music. Their distinction might explain why the man has a sense of misunderstanding in a case where the music being played in the metro station is not familiar.

As previously defined, expectation refers to the cognitive ability of *pre-presenting* a future event that is not well defined, while anticipation is a ‘quasi-perception’ of a highly expected event. When listeners are not familiar with the music they are hearing, they are not able to anticipate the future ongoing of a musical event. For this reason, they have a sense of misunderstanding of that particular music. However, they are able to create dynamic expectations that allow for basic musical understanding and meaningful musical experiences that are subjectively characterized.
CONCLUSION

At the beginning of the paper I proposed the case of the man in the metro in order to introduce some problems concerning the perceptual experience of music that listeners have in complex auditory environments and everyday, sometimes unexpected, situations. As I have tried to show, the act of perceiving music involves various cognitive processes and experiential knowledge. At first thought, all that seems to be needed in order to perceptually understand music is simply to ‘open one’s ears’. However, this notion neglects the complex activity of both the human mind and the auditory system, which is not a passive receiver of information.

According to the phenomenological description of perceptual activity, the human experience is already meaningful because perception allows the immediate awareness of an event or object in the world as having a specific meaning for the subject. Thanks to complex mental processes (such as expectation, anticipation and grouping) and perceptual structures (such as intentionality and temporal structure), meanings are constituted in consciousness without the mediation of conscious reflection. The man in the metro, for example, is immediately able of perceiving music within his complex auditory environment without consciously reflecting on it. In this sense, in order to perceive music and thus understand a sequence of acoustic events as musically meaningful, the first step is, indeed, to ‘open one’s ears’.

Through reference to the case of the man in the metro and by focusing on what he perceptually does in order to be aware of music, I addressed three main questions relevant to both the cognitive psychology of music and phenomenology.

The first question considered the problem of how the awareness of a musical event – a temporally extended object – is constituted by the man among the complex auditory environment of the metro station. I argued for the idea that the man constitutes his musical experience and immediate perceptual awareness of music through a process of ‘listening in
the moment’ – the ability to follow and be involved with music at the same time as listening to it – and the unconscious use of cognitive mechanisms which allow him to organize sounds into musical forms. These cognitive mechanisms may include, for example: similarity, sound localization, recognition of the sound source and learning the regularities of the auditory environment. In this way the man in the metro relates specific acoustic events in a way that allows him to constitute subjective musical meanings in the present moment of the perceptual act, thus developing a basic musical understanding of the auditory environment. The case of timbre-based expectations that I proposed in the seventh chapter aimed to provide an example of how such a process is dynamically realized. The ability to relate sounds in time, which I investigated in the sixth chapter, is therefore essential for acquiring knowledge and comprehension of what is happening around us.

The second question is aimed at investigating the mental processes that determine the subjective character of the man’s musical experience in the metro and his comprehension of the auditory environment. As we have seen, ‘basic musical understanding’ is related to a process of constitution of meaning which is essentially subjective and influenced by one’s own cultural influences, musical knowledge and experiences. The meaning the man’s auditory experience acquires (in the analyzed case, that of being first of all a ‘musical’ experience) thus reflects the way he listens to music and organizes sounds at the perceptual level.

I replied to the second question by focusing on the expectation process, which I consider particularly relevant due to its significant influence on the constitution of meaning. On the basis of the expectations (musical or others) that a listener unconsciously creates, her auditory experience acquires specific musical meanings. In particular, I proposed a conceptual distinction between expectation and anticipation, arguing that they determine different meanings, experiences and bodily responses to music.

The expectations and anticipations (if any) that the man in the metro creates while perceiving the music that someone (one or more musicians) is playing somewhere in the metro influence his perceptual awareness of that auditory event: it might be interpreted as ‘jazz music’, a ‘Christmas song’, a ‘meaningless piece of music’, a ‘sad song’, a ‘guitar’s
version of Leonard Cohen’s *Hallelujah*, or a ‘bad performance of a well known piece of music’.

As previously defined, expectation refers to the cognitive ability of *pre-presenting* a future event which is not well defined, while anticipation is a ‘quasi-perception’ of a highly expected event. When listeners are not familiar with the music they are hearing, they are not able to anticipate the future ongoing of the musical events from one moment to the next. For this reason, they have a sense of misunderstanding of that music. After sufficient exposure and thanks to the process of learning, they will come to understand and anticipate the music they are listening to. These processes, which are essentially unconscious, influence the meaning a listener’s perceptual experience acquires, as well as her responses to music.

The basic idea is that according to most empirical studies, people unconsciously create expectations about the ongoing of music. This is possible because, based upon past experience, they categorize what they are hearing as belonging to a specific musical style. This categorization – or the ability to categorize musical elements - determines the meaning their perceptual experience acquires (which might be emotional, related to compositional features of music or to personal musical experiences).

Giving particular attention to the case of unfamiliar music, the aim of the third question was to provide the basis for an analysis of the man in the metro’s encounter with unfamiliar music. Based on the distinction between expectation and anticipation, I discussed the hypothesis that a lack or incapacity to form anticipations is one of the reasons for the sense of misunderstanding the man might have upon hearing unfamiliar music. However, I suggested that even in this case, he is able to have a kind of basic musical understanding and constitute musical meaning. Timbre organization and timbre-based expectations, for example, may account for his ability to understand an unfamiliar sequence of sound events as *music* and have emotional experiences which influence his appreciation of that music. The structure he perceives likely does not correspond to the notated one (if a notational system exists for the music in question), but he has the capacity to constitute a musical
experience due to the dynamic character of perception and the process of ‘listening in the moment’.

The arguments I presented thus support the idea that a listener dynamically creates expectations as music unfolds even during her first exposure to music or in listening to culturally unfamiliar melodies. In this way, her auditory experience acquires a subjective meaning.

The analysis I have conducted is an attempt to provide new insights for the philosophical investigation into the human experience in the living world and the way people become familiar with the environment. In particular, the discussion considering the way listeners perceptually understand music may suggest new ways of thinking about how human beings understand the world in which they live.

To accomplish this task I focused on some cognitive processes and perceptual structures that are particularly relevant to an immediate perceptual awareness of music and musical understanding of acoustic events occurring in the auditory environment. Focusing particular attention to everyday experience in the world, expectation and anticipation are presented as fundamental principles that contribute to the knowledge of the world and the meaning experienced objects and events acquire for a subject. As we have seen, the ability to create expectation and anticipation influences, in different ways, the comprehension of both familiar and unfamiliar musical events. The analysis of the temporal structure of human consciousness and the concept of ‘listening in the moment’ represent a way to describe how the immediate comprehension and constitution of meaning is developed.

The question concerning the way listeners perceptually understand music and attribute a musical meaning to their auditory experience is common to different disciplines and areas of research. From a musicological perspective a better comprehension of the way music is perceived by an audience might be considered helpful in studying how to use music as an expressive tool that communicates meaning. From a phenomenological point of view such a problem is related to the attempt to describe the way human beings constitute meanings at a perceptual level – how they intentionally form objectivity by virtue of their subjective encounter with the environment. As previously seen, one of the
phenomenologist’s aims is to uncover the essential structures of the intentional consciousness. Husserl, for example, sought to explain the *a priori* structures of the temporal experience as such, through an approach which is radically different from that of the empirical sciences. While the latter try to explain the human consciousness objectively, phenomenology accounts for the subjective aspect of the temporal experience. Despite this difference, cognitive science shares with phenomenology the attempt to understand the human processes of knowledge and the way the human mind works.

Based on this common interest and through an interdisciplinary approach, I endeavoured to create a dialogue between phenomenology and the cognitive psychology of music. In order to go beyond the historical gap between these two fields of research, I think it is firstly necessary to clarify the terminology which is specific to each of them. Some terms are often used in a quite different way and their explanation may help the comprehension of concepts which have a key role in the description of human processes of knowledge.

For example, the analysis of how protention, expectation and anticipation are defined within the phenomenological account of the temporal structure might be helpful in developing the distinction between expectation and anticipation as two slightly different but strictly related cognitive structures which influence the sound information processing and the constitution of meaning at perceptual level. Within the field of cognitive psychology there is often a misunderstanding about what protention is, which results in the frequent use of protention and expectation as synonymous terms. One critique of Husserl’s work is that he did not explain in detail what his definition of expectation was, which is another reason of the misunderstanding surrounding such terms A careful philosophical analysis may contribute to a more thorough definitional approach to these concepts and help to create grounds for fruitful interdisciplinary collaboration.

In addition to the confusion concerning protention, expectation and anticipation, the phenomenological account of the temporal structure of human consciousness is often misunderstood, which makes the dialogue even more difficult. Music is a process in time, characterized by a *succession of impressions* that are related into a same and continuous
(durational) unit. As we have seen, the basic level of the temporal structure is fundamental for the constitution of relationships between sounds and their perception as a unitary whole. Without it connections are not possible. Retention, primal impression and protention allow the creation of continuity and movement between sounds during the temporal unfolding of a musical event. They also guarantee unity to our temporal experience of a whole of sounds. It is because of these modes of appearance that a listener perceives a sequence of sounds as a unity enduring in time.

The musical form itself is centrally a matter of cogency of succession, moment to moment and part to part relations, but the structure it acquires depends on organizational processes that are developed within a temporal movement that involves memory and expectation. While retention, primal impression and protention represent the condition under which to relate sounds in time, memory and expectation determine how we relate them in order to constitute meanings. For example, they allow us to understand why, for a listener, something ends in one way or the other, or why some patterns are used instead of others. Because expectation, anticipation and memory operate on the basis of some cognitive organizational principles (like similarity, grouping and segmentation) contributing in a strong way to the constitution of meaning, and because they themselves are cognitive processes, the cognitive psychology of music contributes to a comprehension of the way sounds are related in time and why they are related in one way or the other. Something that phenomenology does not explain.

The retention-protention structure thus allows ‘listening in the moment’ and guarantees the possibility of having a perceptual musical understanding in hearing a sequence of sounds. The specific meaning (emotional or other) such auditory experience acquires for a subject involves memory, expectation and anticipation. Thanks to both the first and second level of the temporal structure, sounds are intentionally interconnected in order to constitute meaning.

Through the clarification of these concepts we may create a fruitful dialogue between phenomenology and cognitive psychology of music which may allow for a more satisfying
and complete description of the human perceptual experience of music – a description that involves both its subjective and objective aspects.

Another interesting example of how a dialogue between cognitive science and phenomenology may be established on the basis of shared concepts comes from a comparison between the idea of statistical learning and the concept of the horizon. Basically, both approaches explain expectation as referring to the range of possibilities that are unconsciously evaluated by a subject as the most probable to occur among all the possibilities offered by the temporal horizon, which refers to the future. Such probabilities refer to a hierarchy of possibilities which are subjectively built on the basis of memory and acquired knowledge about the most recurrent structures in a particular cultural and social environment.

The probabilistic character of the expectation process is thus acknowledged by both disciplines and it reflects a basic way of knowing and acting in the world. This is another step toward a common and more complete description of the human perceptual knowledge of the environment.

Therefore, the contribution of my dissertation consists of an attempt to clarify some temporal and cognitive structures that are at the basis of human perceptual awareness and comprehension of the world. This may contribute to the understanding of the way human beings know the world, become familiar with it and learn how to react to their multifarious experiences. Further studies are necessary to better understand the human experience in the world and the way an interaction between cognitive psychology and phenomenology may increase such knowledge. In particular, empirical evidence is required in order to clarify some of the suggested hypotheses.

For example, through an empirical investigation about the specificity of expectation and anticipation (what a listener expects or anticipates at any moment when listening to music) we may find evidence that supports this slight but influential distinction. For this reason, I suggest an analysis of the different bodily and emotional responses that people have when they listen to more or less familiar music. Such responses (for instance, changes of posture, metabolism, or conscious thought that prepares the individual for certain
possible outcomes but not for others) might be different in the case where listeners anticipate or simply expect future sound events. For example, expectation prepares the subject to respond to some probable events, while anticipation allows the subject to act in response to the anticipated event before it occurs. Based on the results of such an investigation, it might follow that researchers could pursue an analysis of the influence expectation and anticipation have on listeners’ appreciation of music through a discussion of their aesthetic judgements.

The results of these kinds of studies may have important implications not only for a better comprehension of the human processes of knowledge, but also for the study of people’s social behaviour. Knowing that different responses are related to anticipation or expectation may be useful in particular education programs in order to obtain particular responses and reach some specific results. The difference may also have consequences in business areas and marketing. For example, playing music that people are able to anticipate (i.e. able to sing in advance) in a new bookshop may help them feel better, relax, and become familiar with the new place. Finally, the distinction I have proposed is important from a medical point of view. Music therapy has recently obtained amazing results by using music when treating patients with, for example, brain damage. Knowing that the patient is able to anticipate some music may be helpful in establishing a treatment plan.
REFERENCES


Mondatori.


